

Towards the next trade war?

The trouble brewing on trade between Japan and the United States looks as if it may turn nasty. Both sides are to blame, but the rest of us will be hurt if the worst happens.

THE good news is that there has not been a trade war between the industrialized nations since the ending of the Second World War. The bad news is that the series of disputes bubbling between Japan and the United States over access to each others' markets in advanced electronics could easily become the open war whose avoidance has maintained the remarkable growth of world trade for the past four decades. More constantly than any other, the succession of governments of the United States has played a crucial part in determining this pattern on increasingly open markets. There is no meaningful dispute about the benefits we have all enjoyed from the liberalization of markets during this long period. Industries that were once parochially successful have been able to sell to the whole world, enhancing their prosperity and that of those who work for them. What has been happening is the extension of Adam Smith's doctrine of the efficient division of labour to the international economy.

What has cast such a long shadow over the happy prospect that liberalization would continue indefinitely? There must be clues in two sombre developments reported elsewhere in this issue (p. 319). First, by breathing heavily, Mr Malcolm Baldrige, the US Secretary of Commerce, has managed to dissuade Japanese Fujitsu from going forward with its offer to purchase 80 per cent of Fairchild, a US company well-known as a supplier of semiconductor chips to the US Department of Defense, which has been on the block because its parent company, French-registered Schlumberger Limited, has been badly affected by the slump in oil exploration in the past few years. No legal instruments were invoked to secure Fujitsu's retreat, but only the threat that it might be necessary to do so. Second, in a more subtle because more effective development, Japan has arranged that a consortium whose technical expertise is dominated by foreign companies (including the British-registered company Cable and Wireless) seeking a licence to provide overseas telecommunications services for Japan should be merged with a Japanese latecomer to the competition. The significance of that development is that it weakens Japan's case for claiming that companies elsewhere would be more successful at selling to Japan if they were able to manufacture goods of a quality acceptable to discriminating Japanese customers (which remains, in many cases, the plain truth).

The trouble, so far as it afflicts the United States, is that the manufacture of electronic components, and especially of semiconductor chips, has become a symbol of nationhood. In the language of the business schools, rationality would dictate that manufacturers of equipment incorporating microchips should buy their components from the cheapest supplier, reckoning thereby to increase the margins on its finished products. Two unrelated sets of calculations have muddled this simple picture. First, national security: the snag there is that the argument about the folly of allowing one of the tradesmen to the Pentagon to be owned by a foreign company is undermined by the consideration that Japan is among the most dutiful of the allies of the United States, and that there is no pattern of trade that changes more quickly than that in microchips. Second, and emotionally much more immediate, the United States is convinced that it is being taken for a commercial ride by Japan. What else can be made of

the deficit on overseas trade, which amounted to no less than \$180 million last year? The defect in that argument is the miscalculation that the US trade deficit is a sign of foreign (predominantly, but not exclusively, Japanese) exploitation of the United States. Rather, it is the arithmetical consequence of the US budget deficit. How else than by borrowing from overseas could the United States government consistently outspend its revenue from taxes while refraining from printing money? Last August's microchip agreement between the United States and Japan, doomed to failure from the start for its lack of a convincing definition of fair price (marginal and average costs differ remarkably in capital-intensive industries) was in retrospect an attempt to demonstrate that two small integers do not add up to what childhood lore maintains.

Japan is also not free from blame. Naturally, there is no particular reason why Mr Nakasone's government should have decided three years ago that its trade in telecommunications should be thrown open to all-comers. It could just as well have followed, say, West Germany in the conviction that a nation's telephone conversations with others are too important to be run by anything but an oligopoly of the originators of the messages. Unhappily, for Nakasone's government, it did decide that the time had come to open its telecommunications market. To say the least, it now looks bad that it has closed that market by the old-fashioned subterfuge of pretending that consensus, even at pistol-point, is better than flagrant competition, giving comfort to its critics in the process. And has the famous Japanese consumer had a say?

Bystanders to these parallel but non-intersecting lines of argument, most of Western Europe in particular, must ask themselves two separate questions and then try to find an answer. They, like Japan and the United States, are tainted, in their regard for the principles of free trade and the better division of labour, by their practice of doing almost everything they can to ensure that the downside of compliance will hurt others, not them. The result, in Western Europe, is a common market whose members are as much divided from each other by their petty (but sometimes major) trade barriers as they are united by the conviction that Europe matters. Last week's failure by the European Commission to secure the permission of its political masters to fire a warning shot over the bows of price-fixing airline cartels is relevant. If the world's two most prosperous nations are to have a trade war, should the others band together in neutral mutual self-advantage?

The second more difficult question is that of whether the impending trade row between the United States and Japan can be headed off. If the US presidency were stronger, its successes of the past few years could more easily be headed off. If the US Congress were less frequently compelled to run for re-election, the pressure on congressmen would also be less. But to ask for other circumstances in a pre-election year makes nonsense. The best hope is that the two contenders for the world-title in high technology should call off their contest, and compete for something else instead. Because that will not happen, the next best, but slender, hope is that one or other will recognize the dangers before the fight begins. □



Last hurdle cleared before US-French AIDS accord

- New foundation from royalties of AIDS test
- US-French interests agree draft settlement

Washington & London

THE details of a final settlement to the US-French dispute over the patent rights to a diagnostic test for AIDS (acquired immune deficiency syndrome) have been drafted by the two interested parties and could spawn an international research foundation devoted to the disease. One of the last obstacles to agreement was removed on Sunday, 22 March, during a two-hour meeting at the Intercontinental Hotel in Frankfurt, between Dr Robert Gallo from the US National Cancer Institute and Professor Luc Montagnier of the Pasteur Institute in Paris.

The final version of the agreement is now expected to be signed during the visit by the prime minister of France, M. Jacques Chirac, to the United States on 29 March. Chirac is due to visit President Ronald Reagan at the White House. The French minister of health, Michell Barzach, is in Washington this week for a conference on AIDS.

Gallo and Montagnier have both applied for patents in the United States for the use of antibodies as a basis for testing for infection by human immunodeficiency virus (HIV). This is essentially the test now widely used for assessing the infectious state of people and for deciding whether blood given for transfusion is contaminated.

The French patent was applied for in December 1983. Four months later, the Gallo patent was applied for and was swiftly granted by the US Patent Office, and the then US Secretary of Health and Human Services, Mrs Margaret Heckler, called a news conference to announce a decisive step forward in the then new battle against AIDS. Her enthusiasm added insult to the injury of the neglected French patent application.

The award of the Gallo patent was soon challenged by the Pasteur Institute in what is known technically as an 'interference', the appeals procedure provided by the US Patent Office. The judicial hearings that would normally long since have followed have been postponed on several occasions in the belief that the parties could reach an amicable settlement.

Also in dispute has been the use made by Gallo of a virus isolate provided by Montagnier in September 1983. In accepting this isolate, Gallo and his associates agreed to use it for investigative purposes only, but the French have maintained that the NCI team used the material in the

development of its own blood test.

Under the draft agreement, royalties from the use of HIV antibody tests for AIDS infection will be divided three ways, with one-third of the income going to each of the US Public Health Service (which, through the National Institutes of Health, is ultimately responsible for the National Cancer Institute) and the Pasteur Institute. The remaining third of the income will be directed towards a new international foundation, yet to be established, for the support of AIDS research and treatment, especially in Africa.

At the request of the US parties to the draft agreement, a precondition for the settlement of the dispute has been that there should be an historical statement of the contributions of US and French groups to the discovery and characterization of the AIDS virus. This historical summary takes the form of a chronology going back to the characterization of the virus known as HTLV-I, which is responsible for the adult T-cell leukaemia endemic in southern Japan.

The final form of this document, negotiated in Frankfurt only last Sunday, is believed to include a preamble noting that collaboration between the Gallo and Montagnier groups has never come to a halt, and promising continuing collaboration for the future. An earlier narrative account of the history of research on the AIDS virus was at the outset commissioned from Dr Jonas Salk, but has been replaced by a much more laconic statement. *Nature* hopes to publish the final version of the chronology next week.

The importance attached to this chronology, at least so far as the US side is concerned, is a measure of the degree to which the dispute has been dramatized and even personalized by articles appearing in the popular press, both in France and in English-speaking countries. The letter appearing elsewhere in this issue (p. 326), signed mostly by Nobel prizewinners, seems intended to sanitize this issue.

The main substance of the draft agreement will deal with the apportionment of patent royalties from the application of the antibody tests as they are used at present. The financial importance of the arrangement, and thus the potential of the proposed foundation to make a substantial impact on the problem of AIDS in Africa or elsewhere, will depend on the patents generating income for some substantial time. □

Dispute over plutonium in West Germany

Munich

A ROW has broken out between the West German government and the state of Hesse over the handling of plutonium by the nuclear industry. The Hesse government is refusing to enact a directive from German environment minister Walter Wallmann (Christian Democrat) under which the nuclear fuel manufacturer Alkem would be allowed to process up to 2.5 tonnes of plutonium at a time.

Hesse has instead brought suit in the German Supreme Court at Karlsruhe challenging the federal government's jurisdiction over the plutonium industry. Wallmann is also reported to be considering a counter-suit.

Alkem, based at the city of Hanau, obtains both uranium and plutonium from reprocessing facilities in France. Britain and West Germany and prepares fuel for light-water reactors as well as, potentially, breeder reactors. The present permitted level of 460 kg of plutonium at the plant at one time still allows for several tonnes of plutonium to pass through the plant each year.

Coming as it does just weeks before the 5 April state elections in Hesse, the plutonium dispute has provided fertile political ground for candidates of all parties, including Wallmann himself, who is running for the conservative Christian Democrats. Most striking among the parties' positions is that of the Social Democrats (SPD), who have done an about-face on the issue of nuclear power in the past two years, swinging from a pro-nuclear position to favouring the abandonment of nuclear power in West Germany over the next ten years.

In this, however, they are not as extreme as their coalition partners in Hesse, the Greens, who demand an immediate shutdown of the German nuclear industry. In fact, the so-called "red-green" coalition in Hesse was dissolved in February over the licensing of Alkem.

The contamination of at least 14 workers with plutonium at Alkem's sister company Nukem in February has added to the uproar. But as the fuel containers are said not to have been opened since their manufacture in 1970, it seems unlikely that Nukem will ultimately be blamed for the accident.

The leader of the SPD in Hesse, Ernst Welteke, admitted that the Hessian government's Supreme Court case will probably not get very far. At best, he said, the case may force the Bundestag (the federal parliament) to face the issue of the plutonium industry. **Steven Dickman**

United States fights to protect its semiconductor industry...

Washington

TRADE reprisals against Japan could be announced by the United States within the week in the wake of US frustration at alleged Japanese trade malpractice.

The new measures have been expected after US opposition to the attempted purchase by a Japanese company of the US semiconductor group Fairchild Semiconductor Corporation.

The plan by Fujitsu Limited to buy Fairchild apparently foundered after members of the Reagan administration — notably Commerce Secretary Malcolm Baldrige — expressed their opposition to the takeover. The antagonism toward the Fujitsu deal is symbolic of growing frustration in the United States over Japanese trade practices.

From the US government's perspective, the Fujitsu deal underlined the weakness of the US industry, and reinforced concern about national security with a vital product increasingly controlled by a foreign company. For US industry, the deal was also evidence that Japanese predatory marketing practices — selling semiconductors for less than their cost — was damaging US companies, and making them easy takeover targets.

But Fairchild officials do not see things that way. Fairchild president Donald Brooks said the deal would have added strength to the domestic industry, provided jobs and brought more sales of US-

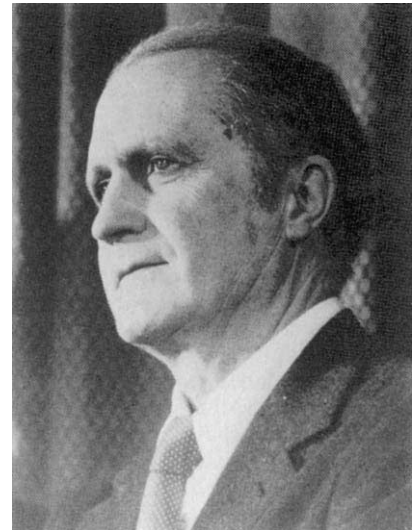
made chips in Japan.

By the end of last week, Fairchild announced it had received several offers of help for a leveraged buyout from its parent company Schlumberger Limited, under the terms of which ownership of Fairchild would have passed to its management. The company still plans to form links with Fujitsu including joint technology development, shared foundry manufacturing and licence exchanges.

Bad timing may have been critical for the Fujitsu-Fairchild takeover. The United States has promised to take action against Japan if violations of a semiconductor trade agreement signed last summer do not stop. That agreement required Japan to stop selling chips below cost on international markets and to open its market to foreign semiconductor sales.

The US industry now claims that Japanese companies have never complied with that agreement. The US government apparently agrees with that assessment and has promised to take some action by 1 April. Action could take the form of duties that would bring the cost of Japanese chips into line with those of chips manufactured in the United States.

But Michael Borrus, of the University of California, says scotching the Fairchild deal and imposing duty does not solve the real problem plaguing the US semiconductor industry. Borrus says Japanese companies have been quietly acquiring



US Commerce Secretary Malcolm Baldrige worries about 'national interest'.

small smart US high-technology companies. This has provided extremely strong prospects for the future, siphoning US technological skill into Japanese industry. Borrus says the Fujitsu deal might have collapsed on its own because of market conditions, but in any case it serves to divert attention from the other more serious issue.

A spokesman for Baldrige says a cabinet-level group is being formed to see if steps need to be taken to protect US companies from foreign takeovers "not in the national interest". But Borrus does not believe that the government is yet taking the problem seriously enough.

Joseph Palca

...and Japan fights to protect telecommunications

Tokyo

AN ATTEMPT to limit foreign participation in a Japanese telecommunications consortium has raised strong protests from the British and US governments and threatens to ignite another trade dispute. At stake is a plan by Britain's Cable and Wireless to lay a network of optical fibre-cables linking Western Europe, the United States and the Far East.

Two consortia have made plans to offer alternative services to Kokusai Den Shin Denwa (KDD) which at present monopolizes Japan's international telecommunications market. One consortium, International Digital Communications Planning Co. (IDC), is 20 per cent owned by Cable and Wireless and proposes to lay its own trans-Pacific submarine cable to link with a Cable and Wireless network that will span North America and the Atlantic Ocean. The other consortium, International Telecom Japan (ITJ), is all-Japanese and intends to lease circuits on Intelsat satellites and a trans-Pacific cable to be laid in 1988 by KDD and American

Telephone and Telegraph.

The Ministry of Post and Telecommunications opposed the IDC plan from the outset. First it was argued that the market is too small for three companies. But even KDD forecasts a tripling in Japan's international telecommunications market to ¥600,000 million by 1995. Then a foreign say in international telecommunications was objected to on grounds of 'national security'.

Now Fumio Watanabe of the Federation of Economic Organizations (Keidanren), with the blessing of the Post and Telecommunications Ministry, has proposed a merger of ITJ and IDC that would slash Cable and Wireless's share to 3 per cent, while key Japanese member companies could take 5 per cent.

With a large number of Japanese shareholder companies in common between IDC and ITJ, there have been no objections from the Japanese side. But Britain and the United States are protesting.

The British Prime Minister, Mrs Margaret Thatcher, in a letter to Japan's

Prime Minister, Mr Yasuhiro Nakasone, says the Japanese moves run "counter to the commitments made by your government about further opening of the Japanese market". And US Trade Representative Clayton Yeutter and Commerce Secretary Malcolm Baldrige have also lodged strong protests with the Japanese government because the US companies Pacific Telesis International and Merrill Lynch are members of the IDC consortium and its cable would land in Alaska.

Jonathan Solomon, Cable and Wireless's director in charge of planning, who arrived in Tokyo last week, protested at the merger proposal, saying "it doesn't make any business sense to invest 3 per cent", and he also urged Watanabe to commit the new group to laying a trans-Pacific cable. But Watanabe says the cable is a matter for the Japanese government to decide and although "Japan is preparing to make its telecommunications market more open, ... the market cannot be opened until we are ready".

David Swinbanks

US Air Force to overhaul its antisatellite weapons programme

Washington

THE Pentagon has reasserted its interest in moving ahead with antisatellite weapons (ASAT) testing by announcing a significant restructuring of its existing ASAT programme. But Congress has prohibited ASAT testing until 1 October this year, and opponents of the ASAT programme moved quickly to introduce legislation that would extend the testing ban.

The last US ASAT test took place on 13 September 1985, when a two-stage rocket tipped with a miniature homing vehicle launched from an F-15 aircraft successfully intercepted an unused satellite in polar orbit. In December 1985, the Air Force launched two instrumented test vehicles (ITVs), intending to use them as targets for additional ASAT tests. But that same month, Congress passed for the first time the Brown-Coughlin amendment prohibiting tests of ASATs for one year so long as the Soviet Union refrained from testing. Congress extended Brown-Coughlin for a second year in October 1986.

The restructuring of ASAT would involve three steps. One would be to test three F-15-launched systems against ITVs during the next fiscal year, with the first test planned for the end of 1987. A second step would be to investigate a more powerful first stage for ASAT launched by the F-15 so that targets in higher orbit could be attacked. At the same time, the Air Force will investigate a ground-launched system using modified Pershing II missiles. A final step would involve development in conjunction with the Strategic Defense Initiative (SDI) of a ground-based high-power laser. For 1988 the Department of Defense (DoD) is seeking

\$450 million for these activities.

The Pentagon maintains that the Soviet Union has had an operational ASAT system since 1982. The Soviet system is said to be ground-launched and to use radar to find its target, destroying it with an on-board bomb. Although DoD is keen to move forward with an ASAT programme, some say this effort is misguided, because DoD is far more dependent on satellites for communications and intelligence than are the Soviet defence services. Many feel an ASAT treaty should at least protect satellites in geosynchronous orbit, as these are critical for electronic surveillance.

Strengthening US ASAT capability is one component of the new DoD Space Policy, signed by Defense Secretary Caspar Weinberger on 4 February. Reaffirming US interest in the military potential of space, the policy statement says that space is "a medium within which the conduct of military operations can take place, just as on land, at sea and in the atmosphere".

The new document replaces one in effect since 1982. One change will be the attention paid to the potential for using men in space to support military objectives. According to Philip Kunsberg, assistant deputy under secretary of defence, "most military functions that we're aware of today can be performed better by machines without man's presence". But Kunsberg asserts that DoD wants to keep its options open for the future.

The DoD space policy specifically endorses the use of both the shuttle and expendable rockets for launching military payloads. DoD has moved aggressively to obtain expendable launch vehicles for its transportation needs, something the advisory council to NASA (National Aeronautics and Space Administration) last week urged that agency to do as well.

The debate over ASAT has been reopened at the same time as a debate over the interpretation of the antiballistic missile (ABM) treaty. Senator Sam Nunn (Democrat, Georgia), chairman of the Armed Services Committee, is leading the congressional push to retain the narrow interpretation he says is justified by the negotiating history and ratification debate. But this would prohibit many planned SDI activities.

The Reagan administration has been considering taking a broader approach to the treaty's provisions, a move that will probably provoke an angry reaction from both Congress and US allies in Europe. The Soviet Union, on the other hand, has proposed an international mechanism for verifying that no weapons are launched into space.

Joseph Palca

East German AIDS

London

THE Institute of Clinical Immunology at East Berlin's Charite hospital has produced monoclonal antibodies against the (HIV) human immunodeficiency virus responsible for AIDS (acquired immune deficiency syndrome). The Charite team see this as a step towards developing antibody-linked drugs to combat the virus. The announcement was made on by Professor Ruediger von Baehr, head of the research team, on East German television.

East Germany has a well-developed AIDS information service and was prepared to discuss the risks to homosexuals and multi-partner heterosexuals at a time when other socialist countries took the official line that such sexual activity was a symptom of 'decadent capitalism'. But the East German media coverage of AIDS is not as explicit as the recent Warsaw television transmission which showed two men in bed together, one removing a condom from its packet.

V.R.

Crafoord prizes 1987

London

THE Crafoord Prize for 1987 will go to Professor Eugene P. Odum of the University of Georgia and Professor Howard T. Odum of the University of Florida, for "fundamental findings which have strongly promoted our understandings of the dynamics of natural systems and formed a scientific basis for the long-term exploitation of the natural resources including pollution abatement", the Royal Swedish Academy of Sciences has announced. The Crafoord Prize, established in 1982, was endowed by the late Holger Crafoord (inventor of the blood dialysis machine) and his wife Anna-Greta Crafoord, and rotates around the branches of science not covered by the Nobel prizes.

V. R.

Mars goal for NASA

Washington

TOP-LEVEL US NASA (National Aeronautics and Space Administration) officials and four independent observers last week discussed future long-range goals for NASA. The previous day the NASA advisory council task force on space goals had recommended that the exploration of Mars be given top priority.

Other goals discussed at the meeting included study of Earth systems from space, an enhanced Solar System exploration and the establishment of a permanent base on the Moon. But none of these was thought to be as attractive as a "bright new vision of Mars".

Although no timetable has been set, contractors will meet on 31 March for a pre-bid briefing on a Mars rover/sample return mission tentatively scheduled for 1998. Meanwhile, Mars Observer remains postponed from a 1990 launch, to 1992.

C.E.

Microbial database

London

A COMPUTER database designed to make microbial culture collections more accessible to industry has been set up at the UK Laboratory of the Government Chemist (LGC). It contains data from the National Collection of Type Cultures, the National Collection of Yeast Cultures, the CAB International Mycological Institute and the National Collections of Industrial and Marine Bacteria.

According to the systems designers, "eventually it will contain data from all the UK National Culture collections enabling subscribers to search from over 30,000 strains. The system is housed on LGC's GEC minicomputer and users can access the information either on-line to the computer or by means of a postal telephone enquiry service.

B. J.

Europe agrees to act for protection of the ozone layer

London

EUROPEAN environment ministers have agreed on the need for international action to control emissions of chlorofluorocarbons (CFCs) into the atmosphere. Under pressure from the United States and from environmental groups, the Council of Ministers of the European Economic Community (EEC) decided in Brussels on 19 March that they are ready to negotiate a global reduction of CFCs.

The EEC will thus attend the conference of the Vienna Convention on the Protection of the Ozone Layer at the end of April with a position in favour of a freeze on production of CFCs 11 and 12 at 1986 levels, a 20 per cent reduction in production of the same materials within five to seven years, a four-yearly scientific review of the evidence on CFCs and restrictions on imports from countries not signing the proposed protocol.

The United States supports a much stronger position, calling for a freeze on production followed by a 50 per cent reduction in five years and a 95 per cent reduction in 10–15 years. The Soviet Union, Canada, West Germany, the Netherlands and the Scandinavian countries are also pushing for big reductions.

Studies by the US Environmental Protection Agency (EPA) show that an 85 per cent cut would be needed just to stabilize current atmospheric concentrations.

Next month's negotiations are aimed at providing the details for a binding international protocol, which is expected to be ready for signing in the autumn by the 28 signatories to the Vienna Convention.

Regulatory measures have been given an additional urgency by new evidence on the Antarctic ozone hole. The results from the ultraviolet spectrometer of the National Ozone Expedition to Antarctica indicate levels of chlorine dioxide 20–50 times higher than those previously measured anywhere, and microwave radiometer measurements of nitrogen oxides and chlorine monoxide provide further support for a chemical explanation of the hole.

There is also a startling new report on the extent and pattern of global ozone depletion. Before a US House of Representatives subcommittee on 9 March, Don Heath, a NASA (National Aeronautics and Space Administration) scientist, reported that observations from NASA's Nimbus 7 satellite indicate a global depletion of total column ozone of several per cent between the mid-1970s and the early 1980s. He also reported a pattern of ozone thinning varying with latitude, with much greater depletion at higher latitudes.

The Heath data are being scrutinized by

a peer-review committee because they suggest latitudinally specific levels of depletion an order of magnitude greater than predicted by two-dimensional models. Such an analysis recently completed for the EPA by Ivar Isaksen of the University of Oslo indicates depletion levels of 0.5 per cent at 80° south latitude, 0.4 per cent at 60° south, 0.4 per cent at 20° and 40° north and 0.9 per cent at 80° north. Heath's data suggest depletion up to ten times as much as these predictions.

Heath also told the subcommittee that his observations point to a second ozone hole over Spitzbergen in the Arctic, al-

though with less severe and extensive depletion than that over the Antarctic.

New information on the human health implications of ozone depletion was also presented to the US Congress earlier this month. Subcommittee members were told that recent analysis by the US government indicates that for every 1 per cent of globally averaged ozone depletion, there will be 10,000–20,000 additional cases of skin cancer a year in the United States alone, with several hundred additional deaths.

Cancer experts say that people who live at 40° latitude or above are considered more at risk for skin cancer, while fair-skinned fair-haired people of Celtic origin have been shown to be the most vulnerable to skin cancer induced by increased ultraviolet radiation.

Kathy Johnston

Another crisis for SERC and British researchers

London

THE Science and Engineering Research Council (SERC), the principal source of funds for British university research, has



Akker and Mitchell — suffering from cuts cancelled its next round of grant awards for lack of cash.

Professor E. W. J. Mitchell, chairman of SERC, had a private meeting last week with the Secretary of State for Education and Science, Mr Kenneth Baker, to highlight the council's plight. The council's poor financial condition has been aggravated by last month's salary award to academic teaching staffs (see *Nature* 325, 567; 1987). The effect in a full year will be to increase SERC's financial commitments by £10 million a year on a budget (for research grants only) of £100 million.

The council has frozen some 1,500 applications for grants until September in the hope that the financial climate will have improved by then.

The ban and the meeting with Baker were the result of a review by the council last week. Commenting on the outcome, Mitchell said that the impact of the pay award on SERC's finances had been so great that, without extra funds from government, the council had "little alternative" but to cancel the April competition for research grants. Mitchell said that discussions with government are con-

tinuing. SERC has, however, "reaffirmed its commitment to postgraduate education" by not reducing the planned number of studentships and fellowships.

John Akker, deputy general secretary of the AUT, says that SERC has "no option" but to take this action. Noting that the volume of civil research and development in the United Kingdom is already desperately low by international standards, Akker says that it is "crazy that our government is causing" the research councils to enforce cuts of this magnitude. According to the AUT, new research projects will now be stillborn and 500 research jobs sacrificed for lack of funds. "The government were party to the pay agreement which has given rise to the problems faced by SERC and the other research councils. The longer the government pro-



crastinate, the worse it will be for morale in the universities."

Bill Johnstone

Northern vision of a research centre for Europe

London

AGAINST the continuing gloom over the state of British research, the launch last week of an appeal to establish a European eye and brain research institute at Manchester provided a gleam of optimism. The sponsors hope to raise £15 million in the next ten years, from industry, commerce and private donors, to provide a self-financing centre for basic and clinical research in the vision sciences.

Professor John Cronly-Dillon, of the department of ophthalmic optics at the University of Manchester Institute of Science and Technology (UMIST), offers as a model for the new centre international laboratories such as the European Molecular Biology Laboratory at Heidelberg. At a meeting of potential backers last week, he declared his ambition that the new centre, to be known as the Northern Eye Institute (NEI), will provide a focus for research into both basic and clinical sciences related to vision and brain function and that it should attract workers from all over the world.

Basic research will range from molecular approaches to genetic and virus-related neurological diseases to machine intelligence. Clinics will be established for research into and treatment of visual complications of diseases such as diabetes

and multiple sclerosis, as well as little-understood problems of visual recognition such as dyslexia. There will be space for industrial researchers to develop commercial applications of instruments and treatments initiated at the institute.

The plan is that the centre should grow in stages, each financed by the interest on the accumulating endowment, with a start on the construction of a building in about five years. Meanwhile, the interest on the invested capital would be used to support vision research in Manchester, with the NEI acting as an umbrella for existing vision research groups at the Manchester University Medical School, UMIST and the Manchester Eye Hospital.

Eventually, according to the plan, the centre should be self-financing on the basis of income from interest, clinic fees, bench fees from visiting workers and overheads on research grants to members of the permanent staff, of which there will be a nucleus on five-year contracts. There will also be income from the expansion of the existing NEI programme of conferences, seminars and teaching courses.

Part of the objective is to find an alternative means of financing research, but there are some fears that initiatives like this could damage the structure of basic research in Britain by absorbing the best

people and by allowing complacency among funding organizations. Those from universities presenting demonstrations of their work at last week's meeting said that vision research in Britain is already falling behind that in Europe and North America. Colin Blakemore, professor of physiology at the University of Oxford (who is in California this week considering two job offers) said that centres such as the planned eye institute should be seen as the pinnacle of a pyramid, not the pyramid itself.

Jennifer Altman

AIDS institute for India?

New Delhi

THE Indian Council of Medical Research (ICMR) is advocating a National Institute for AIDS (acquired immune deficiency syndrome) for research and to treat patients. By 1 March 1987, 104 cases of infection had been detected in India out of 30,000 people so far screened at ICMR's referral centres. At least four new cases are detected every two weeks.

Because many of the infected cases are expected to develop the disease, a hospital exclusively for AIDS is now considered necessary; existing hospitals are overcrowded and are not equipped to deal with AIDS patients.

Indian health officials say that AIDS reached India chiefly through tourists and foreign students. Of the 104 cases so far identified, 11 were foreign students, 11 were tourists and 76 were female prostitutes; only five were Indian males.

Many of the prostitutes in Tamil Nadu claim to have had sexual relations with foreign priests attending Christian conventions. With the discovery of AIDS among Roman Catholic priests, their role in transmitting AIDS is not ruled out, say Indian officials.

Priority for screening has been given to prostitutes because of the large pool of infection among them. But since the Calcutta police gaoled a prostitute suspected of having AIDS, prostitutes all over India have refused to give blood for AIDS tests, which are not at present compulsory because prostitution is illegal. There is a proposal to legalize prostitution and to make the AIDS test mandatory as a condition of granting a licence.

For administrative reasons, the Indian government has dropped the idea of screening tourists, but it has refused to reverse its decision to screen foreign students. Eleven African students have so far been deported. Last week, a Swiss national gaoled for possession of heroin was deported after he said he had AIDS; his accomplice, also a Swiss, has now demanded to be deported by claiming that he shared a bed with his friend in the city gaol.

K.S.Jayaraman

Japan's Ginga (ASTRO-C) X-ray satellite

Tokyo

THE picture shows the launch of Japan's Ginga (ASTRO-C) X-ray satellite at 1500 Japan Standard Time from the Kagoshima Space Center on 5 February (see *Nature* 325, 567; 1987). The photograph was taken by scientists of the Institute of Space and

Astronautical Science (ISAS) with a remote-controlled camera mounted at the top of the assembly tower. Dr Makishima of Tokyo University, who provided the photograph, would like to emphasize that the M-SII-3 rocket is not bent — that is an effect of the fish-eye lens.

The rocket was launched at an angle, rather than vertically, to save power. As the little 28-m solid-fuel rocket streaked heavenward, it was buffeted by strong cross-winds and precious power was used up in course corrections.

As a result, the satellite did not quite make its planned circular orbit of 600 km but entered an elliptical orbit of between 505 and 675 km, a situation which creates a variable X-ray background for the satellite's detectors. But this is not an insurmountable problem according to Makishima.

Within days of the launch, ISAS scientists had something spectacular to test the satellites large-area proportional counters on — the supernova in the Large Magellanic Cloud. But, unfortunately, an X-ray source (LMC X-1) only 0.6 degrees away is obscuring their observations.

David Swinbanks



India seeks scientific basis of traditional remedies

New Delhi

TAKING its cue from China, India has launched a scheme for placing time-honoured traditional remedies on a scientific footing, in collaboration with laboratories in the West.

Indians have been using herbal medicines for centuries without knowing why they are effective. The most promising traditional medicines will now be subjected to Western scrutiny, laboratory research and clinical trials, and those that pass the tests will be earmarked for production.

One drug was released last month by Mr Rajiv Gandhi, the prime minister. Called 'guggulipid', the potent cholesterol-reducing drug has been derived from the extract of 'commiphora mukul', a plant mentioned in the ancient medical system of Ayurveda. Developed by the Central Drug Research Institute and now manufactured by a private drug company, guggulipid is said to be the first major gift of Ayurveda since reserpin, the cardiac drug derived from snakeroot (*Rauwolfia serpentina*). Similarly, a cervical dilator made from seed husks of isubgol (*Plantago ivata*) has completely displaced from the Indian market the laminaria tent, the product of a Swedish company.

Six promising traditional remedies are under Western scrutiny, according to Dr S. Satyavati, deputy chief of the Indian Council of Medical Research (ICMR), who is directing the scheme for resurrecting time-tested herbal cures under the Ayurveda, Unani and Sidha systems. Ayurveda lists 341 plants with medicinal properties and the other two prescribe remedies based on plants as well as inorganic minerals.

India has 275,000 traditional practitioners, and there are 1,600 hospitals and 13,000 dispensaries offering traditional health care. The indigenous systems claim

to have cures for almost any disease, including AIDS (acquired immune deficiency syndrome), but nobody has bothered to evaluate these claims scientifically, says Satyavati. Because modern drug research is a gamble and Indian drug companies lack resources, priority has been given to testing existing herbal cures.

One traditional treatment under evaluation at four ICMR centres is non-surgical treatment of anal fistula using a thread medicated with herbs described in Ayurveda. Called the 'kshaarasootra' technique, it has so far been used to treat more than 2,000 cases. According to Satyavati, patients now demand this treatment in preference to surgery. It has also become popular in Sri Lanka since Sri Lankan surgeons saw a demonstration in New Delhi.

Other traditional remedies under scrutiny by ICMR are those for filariasis, jaundice, urolithiasis, diabetes and bronchial asthma. Also under clinical trial is the herb 'banjari' (*Vicca indica*) used by tribals of Bihar as an abortifacient. An ayurvedic therapy for rheumatoid arthritis has also recently been clinically evaluated under a project sponsored by the World Health Organisation.

The health ministry, on its own, has been spending some US\$100 million a year in promoting research on indigenous medicines. One outcome is a herbal preparation for malaria caused by *Plasmodium vivax*. Eight Indian companies have been licensed to manufacture this drug, which is made available through the ministry's health delivery network. The Unani researchers also claim a cure for vitiligo, white skin patches caused by loss of pigmentation.

A promising anti-cancer drug derived from the plant *Plumbago zylanica* is being evaluated by biochemists at Captain Srinivasamoorthy Drug Research Institute in Madras, as well as at the National Cancer Institute in the United States. The compound called plumbagin has been shown to reduce tumour growth in rats.

The government has encouraged research in indigenous systems of medicine by establishing separate national institutes for Ayurveda and Unani. Research councils have been set up for drugs research, drug standardization and a survey of medicinal plants.

According to the health ministry, "more and more people are taking to traditional systems". Among the reasons are the inherent faith of Indians that herbs have no side-effects, the rampant adulteration of modern drugs sold in drug stores and the rising cost of hospital treatment. At Chittur, in Andhra Pradesh state, a family of traditional healers who set broken bones using a herbal paste attract more patients in a day than the well-equipped orthopaedic wards of government hospitals in any city in India.

K.S.Jayaraman

Japanese plans to sequence human genome

Tokyo

JAPAN'S Science and Technology Agency has taken its first formal step towards setting up a programme for sequencing the human genome.

The agency, in collaboration with Seiko and Fuji Film, has for several years been developing an automatic DNA-sequencing machine under the leadership of Professor Akiyoshi Wada of Tokyo University (see *Nature* 325, 771; 1987), and a move to begin processing long DNA sequences, such as the human genome, has been expected.

On 12 March, the biotechnology subcommittee of the agency's Council for Aeronautics, Electronics and Other Advanced Technologies held a meeting to set up a working group. The subcommittee, under the chairmanship of Professor Zenichi Yoshida of the chemical engineering department of Kyoto University, includes Wada, several other eminent molecular biologists, medical researchers, the director of the Takasaki Research Institute of the Atomic Energy Research Institute and a commentator from Japan's national television network, NHK.

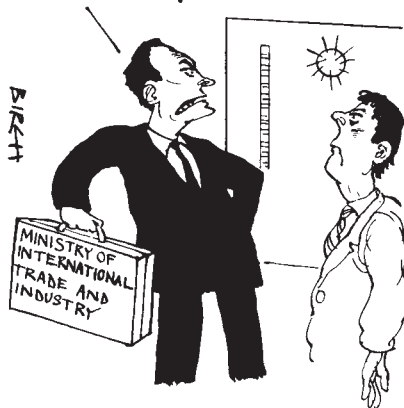
During the next year, the subcommittee will chart policy, and the working group, composed of fifteen scientists including Tasuku Honjo of Kyoto University and Kenichi Matsubara of Osaka University, will draw up a detailed research programme.

Among the issues to be resolved are the relative weights to be placed on mapping and sequencing, and the priority of genes

to be sequenced. Members of the subcommittee and the working group will visit the Seiko and Fuji Film factories to see the automatic DNA-sequencing machine.

There have been suggestions that sequencing of the human genome should form part of the Human Frontiers Science

But there MUST be a gene for Economic Growth!



Program being organized by the Ministry of International Trade and Industry, the Science and Technology Agency and five other ministries (see *Nature* 326, 8; 1987). But Wada, who chairs one of the Frontiers working groups, does not expect genome sequencing to figure prominently in Frontiers, although the development of DNA-sequencing machines may well form part of the programme. The present feasibility study, on the other hand, is devoted specifically to sequencing of the human genome.

David Swinbanks

Freeport for physicists broods about imminent change

Trieste

Is there life after Salam? That is the centrepiece of curiosity at the International Centre for Theoretical Physics here now that it seems certain that Professor Abdus Salam, the founder of the centre and its director since doors first opened in 1964, will next week be nominated as director-general of UNESCO by the government of Italy.

The objective of the centre since its foundation, with funds provided by the International Atomic Energy Agency (IAEA), has been to provide a means by which scientists from developing countries could acquire skills in a field which, however esoteric, has the merit of being portable. Over more than two decades, the still-growing centre, on a promontory with a glorious view towards Venice over the Adriatic, has become known not merely as the Pakistani physicist's creation but as his place. Even when (as last week) he is not here (but in South America), his presence is palpable.

Salam's least predictable achievement is his political popularity in Trieste, once one of Europe's busiest ports which has become an economically depressed city, losing 5,000 young people every year, since the Second World War. Part of the explanation is that Salam's centre has become a model for the attempts by the government of Italy to turn Trieste into a science centre. Italy's first graduate school (known as the International School for Advanced Study) sits alongside the centre. There are substantial centres for economics and geophysics (mostly seismology) already in being. Half of UNIDO's newly created centre for biotechnology will be here (the other half is destined for New Delhi), while it was announced on Monday this week that Trieste has been chosen as the site for Italy's 1.5 GeV synchrotron radiation source.

The funds (roughly \$100 million) for this venture have now been committed by the government. The next year will be spent on a feasibility study, whereafter construction will begin early in 1988. The case for the new synchrotron source, widely supported by the Italian rests on the demands for access to the European synchrotron source at Grenoble.

Whatever the economic motives for these developments, the townspeople believe much of the credit is due to Salam's readiness to make common cause with the city's social and economic problems. One quotes with admiration the once-again prime minister, Sgr Giulio Andreotti, as saying "when somebody from Trieste asks me for something, I say no, but when Salam asks for \$1 million, I

cannot refuse".

Over the years, the government has grown to be the largest single supporter of the centre, now supplying more than 80 per cent of the annual budget of more than \$10 million. The other substantial contributors are IAEA (31 million) and UNESCO (\$400,000 a year so far). There



Professor Abdus Salam — set for UNESCO?

are also small contributions from UNDP and from the government of Britain and the United States (\$10,000 and \$20,000 respectively), the last intended as compensation for the withdrawal of the two governments from UNESCO.

The generosity of the Italian government towards Salam's centre is one of the grounds for the belief that, after Salam, there will be an Italian director. But, others argue, if Salam is snatched away to head UNESCO, he will surely be able to sustain his expressed view that the centre must be run by a physicist of distinction from a developing country. Italian physicists also say that the centre is so different from any other in Italy that only Salam's view could prevail.

The distinctiveness of the centre is beyond dispute. Last year, there were more than 3,000 visitors, most of them theoretical physicists from developing countries, but the remaining people recruited for short spells to run courses or research workshops. The residential full-time staff is a mere handful, but continuity is provided by a group of consultants from nearby institutions, predominantly the theoretical physics department of the University of Trieste (which is housed at the centre). At any time, there may also be between 20 and 30 postdoctoral fellows working at the institute, some of whom will have brought their own funds.

The other main thread of continuity is provided by the group of some 350 associate members of the centre, all of them physicists from developing countries who are elected for periods of six years at a time, and who have a standing invitation to return every other year (on a small

stipend provided by the centre) for periods of between two and four months.

Those concerned speak enthusiastically of this nearly permanent relationship with the centre. Luis Masperi from the Balseiro Institute of the University of Cuyo in Southern Argentina says that almost all the members of the faculty of what is claimed to be the most distinguished institute of its kind in Latin America have been associates of the centre at Trieste. Julian Chela-Flores from the Simon Bolivar University at Caracas, Venezuela, says that the International Institute of Advanced Studies founded there is trying to take a leaf out of Salam's book by providing such a service for the nations of the Caribbean.

One of the most telling arguments in favour of the associateship scheme is that of F. Ardalan, an Iranian visitor to Trieste, who says that in the present political climate, it is virtually impossible for academics from his country (Syria and Libya are in the same case) to go abroad because of visa troubles. Administrators at the centre say that only their good relations with the government of Italy have allowed Trieste to become a kind of freeport for theoretical physicists.

The real bread and butter of Salam's place is, however, the string of courses and workshops which fill the year, and which now cover a much broader scope than the diet of theoretical high-energy physics with which the centre began. Condensed matter physics, astrophysics and an applied physics programme (covering subjects such as management and microprocessors) have been added over the years, not to mention mathematics. One Bangladeshi physicist turned economist said last week that his initial opposition to the inclusion of such an esoteric subject to the centre's field of interest had been dissipated by the success of the programme.

The flies in this Adriatic ointment are few and far between, but they exist. There are some who would like to add further ingredients to the programme — a somewhat tortuous process requiring the approval of the scientific advisory board. The administrative problems of mating the bureaucratic procedures of IAEA with those of the Italian government are more mundane, but more immediate. Who, for example, repairs the building (owned by the University of Trieste as stakeholder for the government)?

Salam has a reputation at the centre for being able to make all these problems go away. The most common worry is that even if he were no further away than Paris, at UNESCO headquarters, the magic would not work. His devotees fall into two groups — those who would like to see him get the UNESCO job and those with a sneaking hope that he is beaten by one of the other candidates.

John Maddox

Stronger links with commerce urged in China

London

CHINESE scientists and technologists must be more prepared to talk business, according to Zeng Xianlin, vice-minister of the State Scientific and Technological Commission. Addressing a national forum of locally based science and technology enterprises in Beijing last month, Zeng said that too many Chinese scientists and technologists are hampered by the "outmoded notion" of looking down on business. Many consider that there are no prospects of engaging in science and technology businesses run by local people. Such erroneous and old-fashioned ideas, he said, should be repudiated.

The movement to create non-governmental locally based science and technology businesses peaked in March 1985, when the Central Committee of the Chinese Communist Party issued a resolution on furthering the reform of the structure of science and technology. Since then, Zeng notes, most research organizations have instituted the technological contract system, and 390 have become economically independent. New technology markets have been opened up and developed, and some 87,000 'technological items' were transferred to the market during 1986. Employment structures have been changed, so that the senior appointments are made on the basis of qualifications and competence, not merely seniority. The most important thing, he says, is that society as a whole now attaches more and more importance to scientific and technological work, and the social status of intellectuals is being upgraded.

Some "defects" in the structure of science and technology still exist, he admits. Scientific and technological personnel are often "dispersed unreasonably", or criticized for doing part-time work in addition to their main jobs. There is still too much bureaucratic interference in the work of research departments, and many research units have not yet established the close relationship with industry recommended by the 1985 reforms. At the beginning of February, the State Council promulgated a number of new regulations on furthering the reform of science and technology. They recommend, on the one hand, the incorporation of independent research and design units into large and medium-sized enterprises; on the other, they encourage the establishment of cooperative relationships between industry and existing research centres and universities and of small technological development agencies operating on a contractual basis.

Vera Rich

Burst raises doubts about Soviet hydroelectricity dam

London

THE dam-burst at the Sargazon reservoir in Tadjik SSR last week with the loss of at least 19 lives has raised some doubts on the future of hydroelectric installations in the republic. In particular, the local *Pravda* correspondent, O. Latifi, has raised the question of the proposed Ziddin reservoir which would mean "millions of cubic metres of water hanging like the sword of Damocles" above Dushanbe, the Tadjik capital.

The designs for the Ziddin project, Latifi complains, were elaborated "without any special publicity". Tadjikistan is a highly seismic area, and the construction of major hydroengineering works requires special consideration. Yet the Vaksh valley has become the site of a massive hydroelectric scheme in which a cascade of eight generating stations will have a combined generating capacity of 9,000 MW and an annual output of some 35,000 KWh.

The two most important stations are the Nurek (designed capacity 2,700 MW, recently upgraded to 3,000 MW) and the Regun, which, with a designed height of 345 m (40 m more than Nurek) will be the highest earthfill dam in the world.

At a recent 'round table' of representatives of the Tadjikistan Academy of Sciences organized at the Dushanbe *Pravda* offices, the scientists expressed their concern about the effect of high dams on the Vaksh valley. A large reservoir is an additional source of seismic stress, and the Nurek dam was built to designs based on the data for plains installations. Yet the effect of a seismic shock of a given magnitude is considerably greater in a mountainous region than in a plain. The torrential rains in the area, moreover, coupled with the low standard of civil

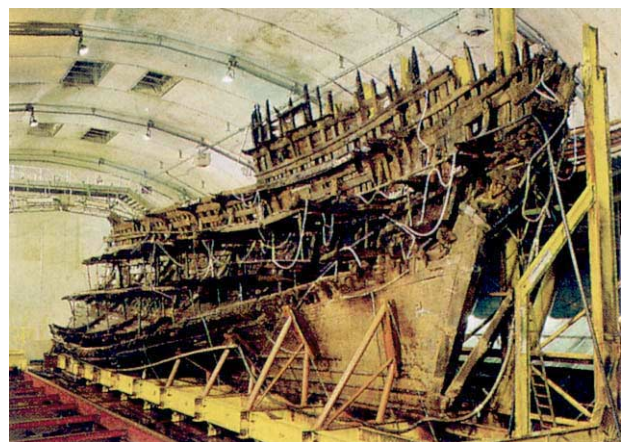
engineering work in the republic, make subsidence and landslips a commonplace. Latifi does not suggest that low standards of work prevail at the Regun high dam, but there were reports last summer that bad coordination between the republic's electrical installation trust (Tadzhikglavenergo) and the All-Union Ministry of Power and Electrification had meant delayed and inadequate supplies, a situation in which pressure to meet target dates (already twice postponed) could easily lead to skimmed or shoddy work. Tajikistan, moreover, like some 25 per cent of the mountain and foothill regions of the Soviet Union, is subject to mudflows. The Sargazon dam-burst was caused by such a mudflow, initiated by torrential rains. Research into mudflows has been going on in the Soviet Union since the 1930s, and there have been several All-Union conferences on the subject. Dams can, in theory, be protected against mud-flows by special mud-collectors, but these do not always prove effective. In July 1973, the impending devastation of Alma-Ata, capital of Kazakhstan was averted only by last-minute action by the State Emergency Commission to pump out the mud-collector at Medeo village. This installation was meant to be the main protection against mudflows for Alma-Ata, but in spite of the (then) more than 25 years of study of such flows, the planners ignored the scientists' recommendations, and design features were omitted.

Latifi's *Pravda* article does not mention the Alma-Ata incident. But something similar, he implies, may lie behind the Sargazon disaster. This tragedy, he says, must "become a lesson". For in it, there was "an element of miscalculation and irresponsibility".

Vera Rich

Mary Rose still under water in dry dock

The Tudor warship *Mary Rose*, sunk on 19 July 1545 and now on public view in dry dock in Portsmouth, is soon to start 'drying out'. The waterlogged timbers have to be sprayed with water for 20 hours in every 24. But now that the *Mary Rose* Trust has conserved all the wooden artefacts from the ship, attention turns to the hull which is to be preserved by a two-stage polyethylene glycol process developed by BP Chemicals.



Hazards of genetic engineering

SIR—The recommendations contained in a recent report of a West German parliamentary commission on applications of 'genetic engineering' (*Nature* 325, 474; 1987) are, in some important respects, disquieting. Accompanying the usual ritual praise of the exploitation of 'genes' and the opportunities for more efficient, economic and ecologically beneficial farming and so on is a recommendation for a five-year moratorium on planned releases of 'genetically transformed' microorganisms. This is illogical and regressive, and, if adopted, would debilitate important areas of basic and applied research in West Germany. Moreover, adoption would send an ominous, misleading message to regulators and others throughout the world that all genetically manipulated microorganisms are hazardous. On both theoretical and experiential grounds, this is emphatically not so.

There are numerous past examples of successful and beneficial 'releases', or uses, of live organisms in the environment. Important examples include the engineering of non-indigenous crops economically important for food and fibre (for example, soybeans in the United States and high-lysine maize in many parts of the world), and the 'creation' of such hybrids as nectarines, triticale and beefaloos. Insect release has been used successfully to control troublesome weeds in Hawaii and California, and the release of a plant pathogen has been used successfully to control skeletonweed in Australia. More than a dozen microbial pesticidal agents are approved and registered with the US Environmental Protection Agency, and these organisms are marketed in dozens of different products for use in agriculture, forestry and insect control. Bacterial preparations containing *Rhizobium*, which allow certain plants to produce nitrogen fertilizer from the air, and which thereby promote the growth of leguminous plants (for example, soybeans, alfalfa and beans) have been sold in the United States since the late nineteenth century.

Other 'deliberate releases' include the use of *Thiobacillus* species, some genetically manipulated, in mining for concentrating certain metals from ores. This has been performed on a vast scale in many parts of the world.

Finally, millions of people each year throughout the world participate, in effect, in 'deliberate release' applications of live attenuated genetically engineered viruses, which undergo replication in the recipient and may be shed. Live human vaccines attenuated by various genetic engineering techniques and now licensed in the United States include mumps, measles, polio and yellow fever. As was

the case with the recombinant DNA (rDNA) pseudorabies veterinary vaccine licensed last year by the US Department of Agriculture, rDNA techniques are expected to improve upon both the safety and the efficacy of some of the products at present produced by older methods. US government agencies and the scientific community generally view the techniques of new biotechnology as extensions or refinements of older techniques for genetic manipulation.

The new biotechnological techniques applied to the production of pharmaceuticals have already yielded genuine medical milestones: interferon for the treatment of hairy-cell leukaemia; a monoclonal antibody preparation to prevent rejection of renal transplants; and sensitive tests for the detection of HTLV-III/LAV antibodies, essential to the safety of the world's blood supply. These successes can be replicated in agricultural and environmental applications, but only if prudent, well-controlled testing is performed.

FRANK E. YOUNG

HENRY I. MILLER

*Food and Drug Administration,
Rockville, Maryland 20857, USA*

Megafauna versus Megafauna

SIR—As a palaeontologist, C. S. Churcher (*Nature* 325, 22; 1987) is intrigued by marine biologists' use of the term megafauna, and raises the question of the appropriateness of the terms megafauna and microfauna, proposes a division of invertebrates into size categories and suggests that "...biologists should get the vertebrate megafaunal and microfaunal categories to overlap those of the invertebrates..."

In marine benthos biology, megafauna is a rather inexact term, mainly used in deep-sea contexts. Operational definitions are "animals readily visible in *in situ* photographs" and "animals sampled by a trawl with part of the bag made of net with a mesh size of 1–2 cm". Some authors do not consider megafauna a separate size category but merely a subdivision of macrofauna, which are animals retained by a sieve with 1-mm mesh size.

Microfauna has been used for several size categories of marine benthic organisms, but most common for those larger than bacteria but passing a sieve with about 50-µm mesh size. The term is now replaced by nanofauna or the more comprehensive nanobenthos.

The size categories in current use for marine animals include both invertebrates and vertebrates. No division according to size exists for the invertebrates alone. To

create one seems inadvisable as there is no need for it and because of the artificiality of the division of metazoa into two groups where one subphylum is opposed to all the rest.

It is desirable that the same prefixes or terms are not used with different meanings in related fields of biology. That situation, however, still exists in marine biology where size categories of plankton and benthos are concerned. A parallel example is the terminology used for terrestrial vertebrate fauna categories and 'biota' of soil biologists. Therefore, to bring marine and terrestrial faunal size categories into line with each other seems a rather remote possibility.

OLE SECHER TENDAL

*Zoologisk Museum, Universitetsparken 15,
DK 1200 København, Denmark*

AIDS agreement

SIR—We are pleased to note the approaching settlement between the United States government and the Pasteur Institute regarding patent rights related to the discovery of the human immunodeficiency virus (HIV), the virus of AIDS (acquired immune deficiency syndrome).

Dr Luc Montagnier and Dr Robert Gallo and their colleagues are the scientists who contributed most to the discovery of HIV and its relation to AIDS. Their work needs to be celebrated and separated from the legal dispute which is being settled through the clarification of the rights and responsibilities of their respective institutions.

It is important to recognize that the discovery of HIV and its relation to AIDS is only the first step towards the ultimate conquest of this disease. We need to encourage our best scientists, both young and older, to engage in solving the urgent problem posed by the spread of this virus in the human population.

DAVID BALTIMORE (*Whitehead Institute for Biomedical Research*), BARUJ BEN-CERRAF (*Harvard University*), PAUL BERG (*Stanford University*), JEAN BERNARD (*French Bioethics Committee*), JEAN DAUSSET (*Collège de France*), RENATO DULBECCO (*Salk Institute*), FRANÇOIS GROS (*Institut Pasteur*), ROBERT HOLLEY (*Salk Institute*), FRANÇOIS JACOB (*Institut Pasteur*), SALVADOR LURIA (*Massachusetts Institute of Technology*), ANDRÉ LWOFF (*formerly of Institut Pasteur*), JONAS SALK (*Salk Institute*), HOWARD TEMIN (*University of Wisconsin*), LEWIS THOMAS (*Memorial Sloan Kettering Cancer Center*), JAMES WATSON (*Cold Spring Harbor Laboratory*), JAMES WYNGAARDEN (*National Institutes of Health*).

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

New ways with matter/antimatter

A computer simulation of the annihilation of particles and antiparticles on a lattice suggests that segregation happens only exceptionally. Cosmologists will be dismayed, but others may be uplifted.

THE notion that some parts of the observable Universe may consist of antimatter keeps cropping up. At an earlier stage, there were suggestions that the great power output of supposedly distant quasars might be explained if they were antimatter galaxies continually being eaten away by the annihilation of their content by interaction with surrounding normal intergalactic material.

Now that it has emerged that there is nothing odd about the radiation from quasars except the amount of it, and in particular no signs of the tell-tale radiation corresponding to electron/positron annihilation, or of that corresponding to nucleon/antinucleon annihilation, the aftermath of the Big Bang has become the most common refuge of the view that unrecognized antimatter may be more common than we know. So it is of some interest that I.W. Anacker and R. Kopelman from the University of Michigan have been able to make a calculation corresponding to a world in which antimatter and matter are being continuously created at a constant rate (*Phys. Rev. Lett.* **58**, 289; 1987). The surprise is that they find one curious state of affairs in which co-existence may be possible.

However solid may be the interests of the cosmologists, it goes without saying that there are many other circumstances in which essentially the same problem arises. The question is to learn a little of the circumstances in which unstable physical systems may exist, perhaps only transiently. Anacker and Kopelman give the separation of electric charges in clouds as one example, but also remind us that Alan Turing, the nearest British equivalent of J. von Neumann in the early days of computers, raised the same question about the existence of complicated biochemical materials a long way from chemical equilibrium. How can these systems apparently continue to exist?

The matter/antimatter problem is the most simply stated. The underlying process is like a chemical reaction between two species whose product, being radiation which plays no further part in the interaction, can be assumed to be literally nothing. So particles of matter (A) and of antimatter ($\bar{A}$) react according to the equation $A + \bar{A} \rightarrow 0$. The ordinary chemical kinetics of this state of affairs are familiar. The reaction is second order. In a well-stirred pot, the reaction rate at any instant is proportional to the product of

the densities of the two species; the density of the least abundant decreases steadily to zero, but strictly speaking will vanish only after an infinite time.

Unfortunately, the Universe is not a well-stirred pot, but rather a place in which particles of matter and of antimatter (if they exist in large numbers) are likely to be physically separate from each other. So the reaction rate will be determined not by the amounts of the two species, or even by their average densities, but by the rate at which matter of one kind diffuses to within annihilation distance of the other. This is where the recondite problem of matter and antimatter has much in common with the more practical problems of heterogeneous catalysis, where the reaction rate is often determined by the speed with which one reactant diffuses to the catalytic surface.

These days, with the advent of supercomputers, the calculation is most simply done numerically. To obtain a non-trivial result, it is best to begin with a system that will sustain itself at least through the duration of a computer run, which is why, having simplified diffusion by the assumption that it consists of random walking on a lattice (one step in a random direction each time unit), Anacker and Kopelman also keep their problem alive by feeding the lattice with equal numbers of particles and antiparticles per unit time. The rules are simply that particles of each kind are added at random to each unoccupied lattice point. In each unit of time, each particle moves one step in some direction chosen at random. If a particle of matter should seek to occupy a lattice point already occupied by a particle of antimatter, or vice versa, annihilation will be the consequence.

People may be forgiven for thinking this an over-simple tale, but the physical unreality of the simulation probably lies more in the assumption that pairs of (A) and ($\bar{A}$) particles appear at places independently selected at random, but in strictly equal numbers, than in anything that may affect the time-evolution of the system. One obvious goal of the simulation is to define the conditions that allow a plausible steady state, defined empirically as a state in which the density reaches some value which is constant in time after 10^6 iterations. Intuitively, it seems clear that, for each lattice, the density of particles in the steady state will increase with the rate at which new particles are

added to the system. The interesting question is whether plausible steady-states accommodate segregation of matter and antimatter.

For a cubic lattice, the steady-state density turns out to be proportional to the square-root of the rate at which new particles are added, but there is no sign of segregation. Segregations are obtained on a linear lattice, but the density keeps increasing with time, which is merely a sign of the inefficiency of random walking on a linear lattice. By the same test, on a cubic lattice random walking is so efficient that the system is always well-mixed.

The surprise in the simulation of Anacker and Kopelman is what happens on the two-dimensional lattice known as the Sierpinski gasket derived from an equilateral triangle by drawing an inverted equilateral triangle within it, and then an appropriately smaller equilateral triangle within each triangular subdivision whose apex points in the same direction as the original. What they discover is that the annihilation process is not incompatible with the segregation of annihilating particles into patches on the computer simulation of the plane. They show a neat two-colour simulation (itself a departure for *Physical Review Letters*) in which there is a concentration of antimatter (or matter, as the case may be) in one-third of the occupied diagram, and only splatterings of it elsewhere.

The moral in all this is anybody's guess, unless it is that the Sierpinski gasket is once again revealed to be a graph (in the technical sense) on which diffusion is hampered by topology; so much has been amply demonstrated in the case of percolation on lattices whose bonds may be represented as pipes of fixed diameter. Certainly the equilibrium concentration of matter/antimatter at which this simulation arrives is too small, in relation to the input of material of both kinds, to lift the spirits of the more radical cosmologists. They, of course, will say that they would prefer to start from a different set of assumptions, allowing matter and antimatter to appear spontaneously not randomly in space, but in a way that allows one ingredient to be systematically separated from the other. Meanwhile, they will no doubt say, there may be some application for this simulation, with all its hours of central processing time on a Cray machine, in chemistry or even sociology.

John Maddox

A quiet supernova?

As astronomers continue to train a variety of instruments on SN 1987a, more differences between it and a generic type II supernova are becoming apparent. In particular, its lack of visibility at some wavelengths may be as significant as its prominence elsewhere.

For radioastronomers, the supernova was almost a dud. An Australian link-up between the Parkes Radio Observatory and NASA's Deep Space Network facility at Canberra detected radio emission on 26 February, three days after the explosion, but by 3 March all was quiet again. This is markedly different from the usual type II supernova which, according to Roger Chevalier of the National Radio Astronomy Observatory in Virginia, may take weeks to warm up in radio frequencies, but then remains bright for several months.

In addition, the Japanese Astro-C (Ginga) X-ray satellite has seen nothing, and astronomers at the Institute of Space and Aeronautical Science in Tokyo deduce an upper limit of 4 milliCrabs on the X-ray

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

flux. (One milliCrab is one thousandth of the measured flux from the Crab nebula.)

Radio and X-ray emission occur as the supernova shock wave ploughs into surrounding gas and dust. The lack of such emission from SN 1987a points to a lack of circumstellar material, and from this argument and others a clearer picture of the so-far unidentified progenitor star begins to emerge.

From the outset, the comparative faintness of the outburst suggests a smaller than usual progenitor, perhaps of eight or ten solar masses rather than twenty. A possibility that comes to Chevalier's mind is of a highly evolved star which had lost a considerable fraction of its mass; a high-velocity stellar wind could both remove the star's atmosphere and blow away surrounding material, leaving a smaller core in an evacuated region of space.

If true, this model has other implications. Infrared emission is likely to be faint, because that too comes from surrounding material 'lit up' by the outburst. And because the fuel supply is small, explosive nucleosynthesis will yield heavy elements only in small quantities, which means that spectroscopy of the expanding supernova shell may be more difficult and less instructive than might have been hoped.

David Lindley

Population dynamics

Fluctuations in a patchy world

John H. Lawton

AN impressive body of ecological theory now points to heterogeneity, be it spatial, temporal or genetic, as an important stabilizing process in population dynamics¹⁻³. It therefore comes as something of a surprise to discover a field experiment, reported on page 388 of this issue⁴, in which increasing spatial heterogeneity apparently destabilizes the interaction between aphids and predatory ladybirds. Peter Kareiva's simple but elegant field experiment involves carving homogeneous stands of the wild plant *Solidago canadensis* (goldenrod) into long, 1-m-wide strips or into more fragmented plots, each 1 m square. Intuitively, many ecologists would probably predict that the more heterogeneous habitat broken into 1-m squares would yield more stable aphid-ladybird population dynamics. In practice, exactly the reverse happens. The more fragmented habitat impedes dispersal by *Coccinella septempunctata*, the predatory coccinellid (ladybird), making it less able to locate and mop up outbreeding foci of its prey, the aphid *Uroleucon nigrotuberculatum*. The result is higher, inherently more unstable aphid populations in the more fragmented patches of goldenrod.

One clear implication of Kareiva's results is that attempts to conserve populations of plants and animals in small, increasingly isolated habitat patches set aside as nature reserves may generate some unexpected and undesirable population dynamics. Outbreaks of herbivores analogous to those created by Kareiva could do serious damage to vegetation and to the entire fabric of the reserve; indeed they may have already happened in some areas⁵. What Kareiva's experiment and related theoretical models⁶⁻⁹ show is that predator and prey mobility between patches of suitable habitat are crucial to the likely outcome of habitat fragmentation. That outbreaks of herbivores are not an inevitable consequence of such treatment is shown by a general tendency for phytophagous insects to be rarer per host plant (that is, to have lower population densities) in smaller habitat patches^{10,11}. Unfortunately, rather little is known about either prey or predator mobility in most of these examples, so the reason why they seem to be at variance with Kareiva's more controlled experiment is unknown.

In general, the *Uroleucon-Coccinella* interaction may be less at variance with existing population-dynamics theory than initially meets the eye because 'stability' needs to be carefully defined¹². Under one definition, stable populations in models or experiments return to equilibrium after a

disturbance. This is an 'all-or-nothing' definition; a population either is or is not stable. Existing theoretical models (for example, refs 2,3) point to heterogeneity as an important contribution to stability in this particular, carefully defined sense. They also show that changing the relative distributions of prey and predators, and the efficiency with which predators seek out prey, alter equilibrium population sizes and can lead either to an increase or to a decrease in victim populations. For example, reducing predator efficiency raises the equilibrium density of the prey², just as Kareiva observes.

A very different use of the word stability (called variability by Pimm¹²) refers to the amplitude of population fluctuations round some long-term average. With this definition, populations can be more or less stable. Moreover, there is no simple relationship between the two, different uses of stability; populations can fluctuate a lot, or hardly at all, and still be formally stable in the first sense of the word. What Kareiva shows in particular is that habitat fragmentation makes aphid populations less stable in the second sense, that is, they show more variability. Theoretical ecologists have paid rather little attention to the factors that contribute to population variability¹².

Kareiva's simple but illuminating experiment also says something about ecology in the broader context of 'hard' versus 'soft' science. There is no shortage of interesting theoretical ideas and sophisticated mathematical models in population biology, many with important practical implications (predicting the effects of habitat fragmentation is but one example). To test his ideas, Kareiva needed nothing more sophisticated than a field of goldenrod, a mowing machine, notebook, hand lens and a great deal of patience and hard work. This is not to say that ecology can always be done on the cheap, but

1. Huffaker, C.B. *Hilgardia* **27**, 343 (1958).
2. Hassell, M.P. *The Dynamics of Arthropod Predator-Prey Systems* (Princeton University Press, 1978).
3. Hassell, M.P. & Anderson, R.M. *J. anim. Ecol.* **53**, 611 (1984).
4. Kareiva, P. *Nature* **326**, 388 (1987).
5. May, R.M. & Beddington, J.R. in *Problems in the Management of Locally Abundant Wild Mammals* (eds Jewell, P.A. & Holt, S.J.) (Academic, New York, 1981).
6. Allen, J.C. *Am. Nat.* **109**, 319 (1975).
7. Hastings, A. *Theor. Populat. Biol.* **12**, 37 (1977).
8. Nachman, G. *J. anim. Ecol.* **56**, 267 (1987).
9. Kareiva, P. & Odell, G.M. *Am. Nat.* (in the press).
10. Strong, D.R., Lawton, J.H. & Southwood, T.R.E. *Insects on Plants: Community Patterns and Mechanisms* (Blackwell, Oxford, 1984).
11. Kareiva, P. in *Variable Plants and Herbivores in Natural and Managed Ecosystems* (eds Denno, R.F. & McClure, M.S.) (Academic, New York, 1982).
12. Pimm, S.L. *Nature* **307**, 321 (1984).
13. Ehrlich, P.R. *The Machinery of Nature* (Simon & Schuster, New York, 1986).

because the subject often has no call for expensive electronic wizardry, it is too often perceived as small, soft science, under the mistaken belief that a million-dollar piece of equipment must somehow be more important than teams of people to count greenfly. Yet more often than not, skilled pairs of hands are what ecologists really need to test and develop theory. The real irony is that because the subject is

still viewed by some as a 'soft' science, we know far more about the stars, the inner workings of atoms and the human body than we do about the fragile, rapidly disappearing and increasingly fragmented ecosystems that are the life-support systems of our planet¹³. □

John H. Lawton is a Professor of Biology in the Department of Biology, University of York, Heslington, York YO1 5DD, UK.

Atomic physics

Single-atom masers and the quantum nature of light

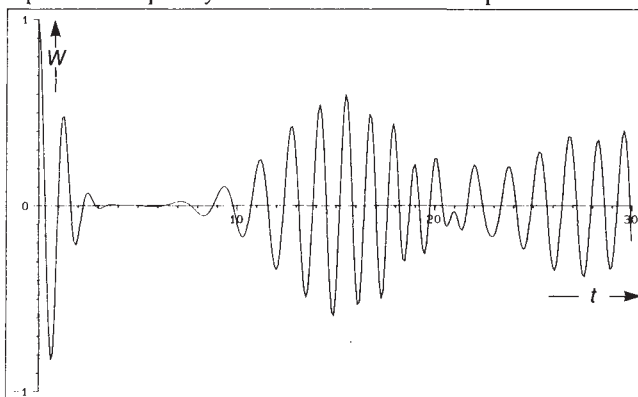
Peter Knight

LASERS and their microwave relatives, masers, usually emit huge numbers of photons, so many that the discrete nature of the radiation is irrelevant to their behaviour. A classical continuously varying field with associated classical noise is perfectly adequate to describe almost all of the ill-named subject of quantum optics.

A dramatic exception is provided by recent work (Meschede, D., Walther, H. & Müller, G. *Phys. Rev. Lett.* **54**, 551; 1985) on masers which comprise only a single atom and radiation fields so well that only two or three photons are present. This 'micro-maser' can measure periodic energy exchange as a single photon is absorbed and re-emitted repeatedly in the maser cavity. Now, G. Rempe, H. Walther and N. Klein (*Phys. Rev. Lett.* **58**, 353; 1987) report observations of the influence of photon statistics on Rydberg atoms in a cavity with a photon storage time so long that they can detect single-photon events. The authors detect quantum collapses and revivals of the maser coherent evolution, long believed to be a subject of interest only to theorists. These collapses and revivals are a fundamental quantum effect that depends entirely on the discrete nature of the photons in the maser cavity.

An atom irradiated by a classical radiation field whose frequency is tuned precisely to an atomic transition frequency is excited to a higher energy level by stimulated absorption. Having achieved the excited state, the atom can be induced to return to the ground state by stimulated emission if the incident intensity of the radiation is sufficiently high, otherwise the atom emits its stored energy incoherently by spontaneous emission. The rates for these emissions and absorption were calculated in 1917 by Einstein. But if the radiation is sufficiently coherent, this simple rate picture is inadequate to describe

the transition dynamics. Instead, the atom periodically and coherently exchanges energy with the radiation in a sinusoidal oscillation first studied by Rabi. Rabi oscillations are one of the basic concepts of quantum optics, and underlie present ideas about resonant excitation. The frequency of the Rabi oscillations depends



Time dependence of the excitation probability, W , of the initially excited atom, driven by an initially coherent field with $n = 5$. The Rabi oscillations are dephased by quantum fluctuations of the field, but partially revive later as the discrete nature of the field becomes evident.

on the magnitude of the electric field of the radiation.

When the radiation field is described by the laws of quantum mechanics, the continuously variable electric field is replaced by the discrete number of photons occupying that mode of the radiation field. A coherent field, rather than having a definite number of photons, has instead a Poisson distribution around a mean n with a width given by $\sqrt{n}$. Normally the photon number n is huge in laser physics and the fractional uncertainty in n is correspondingly unimportant. But an atom can, in principle, detect this uncertainty. Were there to be precisely m photons of a particular frequency interacting with an atom, the atom and field again exchange energy sinusoidally according to the ideas of Rabi.

The quantum Rabi frequency, first studied by E. T. Jaynes and F. W. Cum-

ings in 1963 (*Proc. IEEE* **51**, 89; 1963), is analogous to the classical version, but has the electric field replaced by a quantity depending on the square root of the photon number m . States of the radiation field having precise numbers of photons are unknown. Instead we have to average over the probability distribution for there to be m photons in the interaction region. This leads to a collapse of the Rabi oscillations: the different possible photon numbers lead to a distribution of possible Rabi frequencies that dephase the coherent evolution. The dephasing is not irreversible. Because the photon numbers are discrete variables, the Rabi oscillations will rephase periodically. These revivals are the signature of field quantization (see figure).

To see Jaynes-Cummings collapses and revivals, it is essential to have an atomic transition that couples very strongly to the radiation. Rempe *et al.* used very highly excited Rydberg atoms with transition wavelengths in the sub-millimetre range. These atoms have dipole moments five orders of magnitude larger than normal microwave transitions (for example in the ammonia maser). The Rydberg atoms are produced by laser excitation of a beam of velocity-selected rubidium atoms. These highly excited atoms then enter a superconducting microwave cavity which supports a radiation field capable of resonantly exciting transitions to nearby Rydberg levels. Transitions are detected by ionizing the atoms as they leave the cavity in a static electric field: the number of ions detected depends on which atomic state is populated. Because the atoms are velocity-selected, the precise time available for coherent interaction with the field in the cavity is known. The probability of the atom remaining in its excited state

is found to evolve exactly as predicted by the Jaynes-Cummings model. The sinusoidal Rabi oscillations at a frequency governed by the mean photon number are observed to collapse because of the distribution of Rabi frequencies. In the experiment of Rempe *et al.* the collapse was facilitated by a background thermal field with a broad distribution of photon numbers present in the maser cavity even at the temperature of the liquid-helium-cooled apparatus. At lower temperatures the Rabi oscillations persist for longer and a revival was observed. In these experiments there are on average only 2.5 photons in the cavity. That such a weak field is capable of being detected at all is a tribute to the skill of the micromaser experimentalists. □

Peter Knight is in the Blackett Laboratory, Imperial College of Science and Technology, London SW7 2BZ, UK.

Tumour necrosis factor

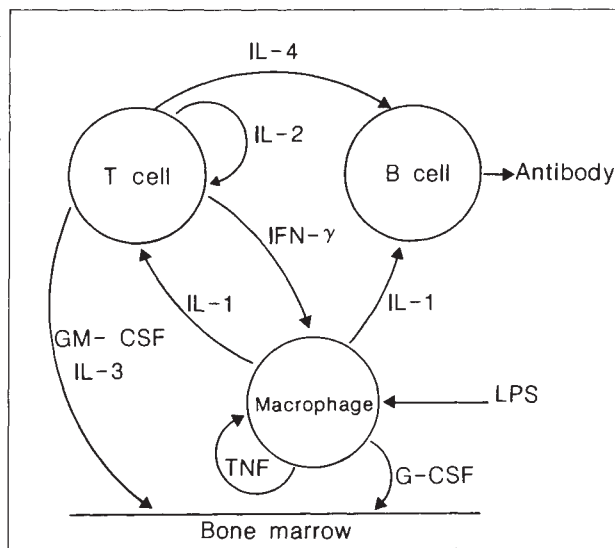
Polypeptide mediator network

Lloyd J. Old

THE list of polypeptides regulating the growth, differentiation and function of cells involved in immunity, inflammation and haematopoiesis has grown rapidly during the past decade. Cloning of these factors provided formal evidence for their existence and distinction from other factors, and the availability of cloned products has greatly facilitated examination of their biological activities. Tumour necrosis factor (TNF) and lymphotoxin (LT), two factors initially recognized because of their cytotoxic and anti-tumour properties, were the focus of a recent meeting*. The dramatic effect of TNF on experimental tumours (haemorrhagic necrosis and regression) has inspired intense interest in its clinical application, but initial trials in cancer patients have not shown similar anti-tumour effects. A broad spectrum of additional TNF activities has been identified, however, ranging from regulatory effects on normal cell types such as neutrophils, fibroblasts and endothelial cells, to inhibitory effects on parasites and viruses.

Although the amino-acid sequences of TNF and LT have only limited homology, they bind to the same receptor, mediate the same responses and map to the major histocompatibility complex region in mouse and man (B. Aggarwal, Genentech). The cellular origins of TNF and LT are different: macrophages are the major source of TNF, whereas T cells produce LT (N. Ruddle, Yale University). Because cloned TNF has been more available than LT, biological characterization of TNF is more advanced. Lipopolysaccharide (LPS), the cell-wall component of gram-negative bacteria, is a potent stimulus for TNF production by macrophages (G. Gifford, University of Florida College of Medicine; K. Haranaka, University of Tokyo). Many of the activities of TNF mimic the well-known effects of LPS, for example, tumour haemorrhagic necrosis, fever, shock and activation of neutrophils, indicating that TNF is an important mediator of LPS action, and antibodies against TNF have been reported to inhibit LPS-induced tumour haemorrhagic necrosis and toxicity. A surprise emerging from studies in several laboratories is the remarkably

similar behaviour of TNF and interleukin-1, another LPS-induced macrophage product with no sequence similarity to TNF or LT and a distinct receptor system. Interleukin-1, formerly called endogenous pyrogen, leukocyte endogenous mediator or lymphocyte-activating factor, also exists in two forms (IL-1 α and IL-1 β), with identical activities and the same receptor, but little sequence similarity. Like LPS,



Polypeptide mediator circuits. LPS and other microbial products initiate release of TNF, interleukin-1 (IL-1) and granulocyte colony-stimulating factor (G-CSF, a haematopoietic growth factor) from macrophages. TNF by itself is also a signal for macrophage IL-1 production. IL-1 activates T cells (either directly or indirectly through accessory cells), and activated T cells produce IL-2 (a T-cell growth factor), granulocyte-macrophage CSF (GM-CSF) and IL-3 (haematopoietic growth factors), IL-4 (a B-cell growth factor) and interferon- γ (IFN- γ). IFN- γ activates macrophages and amplifies LPS-induced TNF release by macrophages. TNF also stimulates IL-1 and GM-CSF production by endothelial cells and fibroblasts. Insofar as tested, LT (a T-cell product) has the same mediator-inducing activities as TNF.

TNF is a strong signal for interleukin-1 production *in vitro* and *in vivo* (M. Palladino, Genentech). Despite the broad range of shared activities, there are several differences between TNF and interleukin-1: first, TNF has more potent tumour necrotizing activity *in vivo* and greater cytotoxic activity against tumour cells *in vitro* (Palladino; myself and co-workers); second, TNF augments major histocompatibility complex antigen expression in endothelial cells, whereas interleukin-1 does not (J. Pober, Harvard Medical School); third, interleukin-1 is produced by a much wider range of cell types than TNF; and, finally, interleukin-1 is an important signal for T-cell activation, whereas TNF is not.

Macrophages activated by various stimuli, including LPS, can inhibit malignant cells to a greater degree than normal cells. Three lines of evidence now directly implicate TNF in macrophage killing: target cells rendered resistant to macrophages are resistant to TNF; target cells rendered resistant to TNF are resistant to macrophages; and TNF antibody inhibits target-cell killing by macrophages (H. Schreiber, University of Chicago). The basis for the cytotoxic/cytostatic effects of TNF and LT on sensitive tumour cell lines is not understood; involvement of prostaglandins, proteases and free radicals, labilization of lysosomal enzymes and DNA fragmentation have been suggested.

Whatever the mechanism, interferons, certain metabolic inhibitors and heat greatly potentiate TNF cytotoxicity; dexamethasone and phospholipase inhibitors (W. Fiers, State University of Ghent) and certain protease inhibitors (C. Baglioni, State University of New York, Albany) block it. The basis for TNF-induced tumour haemorrhagic necrosis is also unclear, but clues about the mechanism are emerging from studies on the direct effect of TNF on endothelial cells (Haranaka; Pober). TNF alters the growth and morphology of endothelial cells, increases synthesis of procoagulant activity and other factors that favour blood clotting, and enhances endothelial cell adhesiveness for inflammatory cells.

Much discussion centred on the role of TNF in the pathogenesis of cachexia associated with infectious or neoplastic diseases. The suggested link with cachexia comes from studies on a factor called cachectin, produced by LPS-activated macrophages, that inhibits lipoprotein lipase, an enzyme involved in lipid storage in adipose tissues (A. Cerami, Rockefeller University). Cachectin and TNF are now known to be identical. Recent work shows that interferons and interleukin-1 also inhibit lipoprotein lipase synthesis. In addition, cloned TNF, although toxic in high doses, does not cause a cachectic state in mice or rats on repeated injection of lower doses, nor has cachexia been seen in patients treated with TNF. Cachexia is a complex metabolic state involving changes in protein as well as lipid metabolism, and TNF does not cause increased protein breakdown and prostaglandin synthesis in muscle as do LPS and bacterial infection (A. Goldberg, Harvard Medical School). Although TNF does not have direct inhibitory effects against microbes and parasites, it protects against some types of malarial

* Tumour Necrosis Factor and Related Cytotoxins Ciba Foundation Symposium No. 131, London 20-22 January 1987.

infections in mice (J. Playfair, Middlesex Hospital Medical School, London; V. Nussenzweig, New York University). In addition, a macrophage-derived factor that activates eosinophils to kill schistosomes in the presence of antibodies has been identified as TNF (J. David, Harvard School of Public Health), and there are elevated TNF levels in the blood of patients with visceral leishmaniasis and malaria. Antiviral activity of TNF against RNA and DNA viruses has also been reported recently, with evidence that TNF has selective cytotoxicity for virally infected cells (Aggarwal). It is not known whether this antiviral effect is direct or is mediated through another factor induced by TNF (interferon- β).

Several phase I clinical trials of TNF in cancer patients have been started (D. Spriggs, Dana-Farber Cancer Institute). Fever, chills, malaise and hypotension are the major side effects, and are similar to reactions in patients treated with LPS, interferons or interleukin-2. Anti-cancer effects have been reported with systemic TNF treatment, but these have been rare in the phase I trials and phase II studies are needed to determine the response rate. There is considerable interest in

combined treatment with TNF and interferon, which arises from the synergistic cytotoxic action of TNF and interferon on cancer cells *in vitro*, and from therapy studies in mice, particularly the impressive therapeutic efficacy of TNF and human interferon- γ against human ovarian cancer xenografts in nu/nu mice (F. Balkwill, Imperial Cancer Research Fund Laboratories).

The general consensus was that TNF is part of a network of interactive signals that orchestrate inflammatory and immunological events and maintain a proper balance of haematopoietic stem cells and their descendants (see figure). Although the role of each component can be studied in defined *in vitro* systems, it is more difficult to interpret the *in vivo* effects after injection of individual mediators because of the complex circuitry involved. Clearly, the rational use of TNF or its antagonists as therapeutic agents depends on a better understanding of its physiological role and its importance in the pathogenesis of infectious, degenerative and neoplastic diseases. □

Lloyd J. Old is at the Memorial Sloan-Kettering Cancer Center, 1275 New York Avenue, New York, New York 10021, USA.

Mass extinctions

Iridium anomalous no longer?

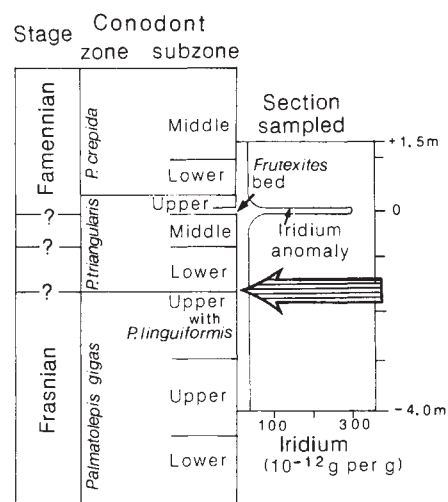
Stephen K. Donovan

IT HAS been suggested¹ that the process of mass extinction of species is extraterrestrial in origin, probably caused by the impact of large meteorites or comets (bolides) on the Earth. The main evidence for bolide impact is enrichment of the clay at the Cretaceous-Tertiary boundary, a time of great faunal turnover, with platinum group metals, particularly iridium¹. The Earth's crust is depleted in these elements relative to their cosmic abundance and it is believed that their anomalously high concentration at the Cretaceous-Tertiary boundary indicates a major non-terrestrial source. But new searches for evidence for the impact of extraterrestrial objects which coincide with mass-extinction events have failed to find support for these theories at three important levels of the geological record²⁻⁵.

Since the Cretaceous-Tertiary anomaly was first discovered, it has become fashionable for geochemists to play 'hunt the bolide'. First, choose your mass extinction (the principal mass extinctions recognized are at the following geological boundaries: Cambrian-Ordovician, Ordovician-Silurian, Frasnian-Famennian (Upper Devonian), Permian-Triassic, Triassic-Jurassic, Cretaceous-Tertiary and Eocene-Oligocene). Then find an iridium anomaly and you have an impact.

This approach has confirmed the anomaly at the Cretaceous-Tertiary boundary. It has also demonstrated anomalies at the Permian-Triassic boundary in China⁶ and the Frasnian-Famennian boundary in Australia⁷. Not every pursuit of an anomaly has discovered iridium in association with mass extinction, however, and even where present, iridium is turning out to be a weak argument for a bolide at some extinction horizons.

Two recent papers^{2,3} report iridium anomalies from the Ordovician-Silurian boundary, coinciding with the end-Ordovician extinction. The traditional view⁸ of this extinction is that it was controlled by glacio-eustatic events which led to the demise of many organisms that inhabited shallow shelf seas. Different groups of organisms were affected at slightly different times, so that this event does not seem to have been a single catastrophic blow. There is widespread evidence for a late Ordovician glaciation in the Southern Hemisphere extending to 40 °S (ref. 8). Despite the discovery of iridium in two Ordovician-Silurian localities, the glacial theory still seems the most probable explanation of the extinction. Orth *et al.* found² an iridium anomaly at the boundary on Anticosti Island, Quebec, an important Ordovician-



The Frasnian-Famennian iridium anomaly in relation to conodont zones and stages (redrawn from ref. 9). The shaded arrow indicates the horizon of mass extinction described in ref. 3.

Silurian sequence, of 58 parts per trillion, about 11 times the minimum of the rest of the section. This is not evidence of a bolide, but is merely a reflection of a high clay content of the sediments at the boundary. Iridium and most other trace elements in the section are concentrated in alumina-rich clays. In the more calcium carbonate-rich rocks iridium is found to be proportionally scarcer.

Wilde *et al.* examined³ the iridium abundance across the Ordovician-Silurian stratotype at Dob's Linn, Scotland. Iridium concentrations lie in the range 0.050 to 0.250 parts per 10⁶ (p.p.b.) through a rock thickness, from the late Ordovician into the early Silurian, equivalent to 20 million years. There is no iridium peak at the boundary. Rather, the amount of iridium at every horizon sampled is unusually high. This iridium was probably derived from erosion of unroofed upper-mantle ultramafic rocks of the Ballantrae Ophiolite at the time of deposition of the Dob's Linn shales. Upper-mantle rocks are enriched in iridium compared with the crust, and platinum-group elements from the ophiolite were presumably weathered and then redeposited as part of the Dob's Linn sequence.

Although a Frasnian-Famennian anomaly has been found⁷ in the Upper Devonian of Australia, it has not been noted elsewhere, despite the global extent of the associated extinction event. In Europe the Frasnian-Famennian boundary corresponds to a widespread sedimentary deposit, the Kellwasser black shales and bituminous limestones. McGhee *et al.* have sampled⁴ this unit in search of iridium. There were two horizons to be considered (see figure). The level of the Frasnian-Famennian extinction event itself does not correspond to the (younger) horizon in Australia that has given an iridium 'spike'. Thus, the Australian iridium anomaly postdates the mass ex-

tion in Europe. The results from the Kellwasser are entirely negative; there is no iridium anomaly, nor is there any other evidence to indicate bolide impact. From the stratigraphical and geochemical evidence of the Kellwasser, it must be concluded that the Australian anomaly postdates the actual mass extinction and therefore cannot be the cause of massive faunal turnover. This iridium 'spike' is probably caused by terrestrial rather than by extraterrestrial causes.

The only report to date of an anomaly from the Permian-Triassic boundary is from China⁶. A further geochemical investigation⁵ of a Permian-Triassic section elsewhere in China has failed to support the first report. Indeed, in the latest section to be investigated, iridium at the boundary is particularly rare, 0.002 p.p.b. ± 20 per cent, compared with the range of 0.004 – 0.034 p.p.b. elsewhere in the sequence. Once again, the possibility of an extraterrestrial impact at a major mass extinction is not supported by the detailed geochemical evidence of more than one section. The earlier report of an anomaly may yet turn out to be an artefact of the sedimentary record.

The sections that fail to support extraterrestrial causes for mass extinction all have excellent biostratigraphical controls, either from graptolites or conodonts. Dob's Linn is the Ordovician-Silurian boundary stratotype; Anticosti Island was a candidate for the same distinction; conodont zonation within the Kellwasser is good; and the new section in south China is regarded as "... one of the most complete and fossiliferous Palaeozoic-Mesozoic boundaries known"⁵. There is, therefore, no doubt that actual boundary sequences have been sampled. The iridium in the two Ordovician-Silurian sections^{2,3} is regarded as being terrestrial in origin. Perhaps similar processes can explain the regional variation in iridium at the Frasnian-Famennian and Permian-Triassic boundaries. At present, however, there is still a very large body of data to support an extraterrestrial influence on the Cretaceous-Tertiary mass extinction. It may be the only extinction that cannot be attributed to earthly causes. □

1. Alvarez, L.W., Alvarez, W., Asaro, F. & Michel, H.V. *Science* **208**, 1095-1108 (1980).
2. Orth, C.J., Gilmore, J.S., Quintana, L.R. & Sheehan, P.M. *Geology* **14**, 433-436 (1986).
3. Wilde, P. *et al.* *Science* **233**, 339-341 (1986).
4. McGhee, G.R. Jr, Orth, C.J., Quintana, L.R., Gilmore, J.S. & Olsen, E.J. *Geology* **14**, 776-779 (1986).
5. Clark, D.L., Wang Cheng-Yuan, Orth, C.J. & Gilmore, J.S. *Science* **233**, 984-986 (1986).
6. Xu Dao-Yi *et al.* *Nature* **314**, 154-156 (1985).
7. Playford, P.E., McLaren, D.J., Orth, C.J., Gilmore, J.S. & Goodfellow, W.D. *Science* **226**, 437-439 (1984).
8. Branchley, P.J. & Newall, G. in *Aspects of the Ordovician System* (ed. Bruton, D.L.) *Paleont. Contr. Univ. Oslo* **295**, 65-79 (1984).
9. McLaren, D.J. *Geology* **13**, 170-172 (1985).

Stephen K. Donovan is in the Department of Geology, University of the West Indies, Mona, Kingston 7, Jamaica.

Molecular evolution

Eukaryotes with no mitochondria

T. Cavalier-Smith

It is a widespread fallacy that mitochondria are found in all eukaryote cells, and that the symbiotic theory of the origin of mitochondria from purple photosynthetic bacteria, for which the evidence is now exceedingly good, explains the origin of the eukaryote cell itself. In fact it is not mitochondria, but the nucleus, endomembrane system and cytoskeleton that are the true hallmarks of the eukaryote cell: there are more than a thousand species of protozoa and a few fungi that have no mitochondria at all. The idea that some of these protozoa are living relics of the earliest phase of eukaryote cell evolution and diverged from our own ancestors before the symbiotic origin of mitochondria^{1,2} is given strong support by DNA sequence studies of one phylum of amitochondrial protozoa, the microsporidia^{3,4}, reported by Vossbrinck *et al.* on page 411 of this issue⁴.

All microsporidia are intracellular parasites either of animals (now also turning up in AIDS patients) or of protozoa that do have mitochondria, and this fact encouraged the belief that the absence from their cells of organelles such as mitochondria, peroxisomes and a typical Golgi dictyosome was not primitive but was a secondary reduction caused by their parasitic lifestyle, for which they have a unique tubular cell-penetration machinery (Fig. 1). But mitochondria are present in all other intracellular protozoa, such as the malaria parasite, and there is no evidence that any intracellular parasites have ever lost mitochondria, although such loss has

undoubtedly occurred in certain ciliate (infusorian) protozoa and spizellomycete fungi, which inhabit the highly anaerobic stomachs of ruminants. The latter conclusion is based on the fact that both these groups also contain relatives that possess mitochondria, which is not so for the four protozoan phyla listed in the table, recently classified (pages 1027-1034 of ref. 1) in the protozoan subkingdom Archezoa because of this and other unique features. Those protozoa which always, or usually (Infusoria), have mitochondria are placed⁵ in the subkingdom Mitozoa.

One of the most important recent developments in molecular phylogenetics has been the use of ribosomal (r) RNA sequences for constructing phylogenies, pioneered by Woese and collaborators. New, rapid DNA sequencing methods using primers complementary to conserved sections of cloned small rRNA genes are yielding a forest of new evolutionary trees. The sequence reported in this issue⁴ by Woese's group for the microsporidian small rRNA gene is the first for any archezoan phylum, and diverges much more from all previously studied eukaryote sequences than any of them do from each other. This is exactly as expected if microsporidia are indeed primitively without mitochondria. That the microsporidia are very primitive eukaryotes is also evident in that, like bacteria, they have small ribosomes (70S) containing 16S and 23S rRNA³, rather than the larger 80S ribosomes, which contain the larger 18S and 28S rRNA molecules

Protozoan phyla without mitochondria

	Cilia per cell	Ribosome type	Golgi dictyosomes	Peroxisomes	Hydrogenosomes
1. Archamoebae					
class Mastigamoebae	1	?	—	—	—
class Pelobiontea (for example, <i>Pelomyxa</i> , <i>Entamoeba</i>)	0	?	—	—	—
2. Metamonada					
class Anaxostylea (diplomonads and retortomonads)	1-8	?	—	—	—
class Axostylaria (oxymonads)	4	?	—	*	*
3. Microsporidia (Microspora)					
class Rudimicrosporea	0	?	**	—	—
class Microsporea	0	70S	**	—	—
4. Parabasalia					
class Trichomonadea	0-5	?	+	—	+
class Hypermastigae	many	?	+	—	+

* Microbodies of some sort are visible in the electron microscope but it is not known whether they are peroxisomes, hydrogenosomes or neither.

** Unlike phyla 1 and 2, presumed Golgi membranes have been identified but they do not have the form of dictyosomes.

characteristic of higher eukaryotes. As in bacteria, the 23S rRNA of microsporidia contains sequences complementary to 5.8S rRNA, which in higher eukaryotes is cleaved away from 28S rRNA to form a separate molecule during RNA processing. Many molecular evolutionists have assumed that the special features of 80S ribosomes reflect an early evolutionary divergence that preceded the origin of the uniquely eukaryotic cell organelles such as the nucleus and cytoskeleton. It is now clear that, on the contrary, 80S ribosomes evolved after the origin of eukaryotes, and also after sex and meiosis, both present in microsporidia. Studies of all protozoan groups are now needed to show whether 80S ribosomes evolved before, after or during the origin of mitochondria, and to exclude the possibility — very unlikely in my judgement — that any mitozoa are more primitive than the microsporidia.

Because of their obligately endopara-

candidates for the host that engulfed purple bacteria by phagocytosis, followed by breakage of the food vacuole, to form mitochondria, and which also engulfed cyanobacteria to form chloroplasts. Contrary to the serial endosymbiosis theory⁶, in which mitochondria were supposedly formed before nuclei and chloroplasts long afterwards, the ultrastructural phylogenetic evidence indicates that mitochondria and chloroplasts originated at about the same time^{1,2,5}, distinctly after the origin of the nucleus. The rRNA sequencing data also support this simultaneous endosymbiosis theory, for the earliest divergence within the subkingdom Mitozoa is between Euglenozoa (the phylum comprising mainly the euglenoids and kinetoplastids such as trypanosomes) on the one hand, and all the other mitochondria-bearing phyla on the other⁴. As both sides of this early split contain species with chloroplasts, this dates the origin of

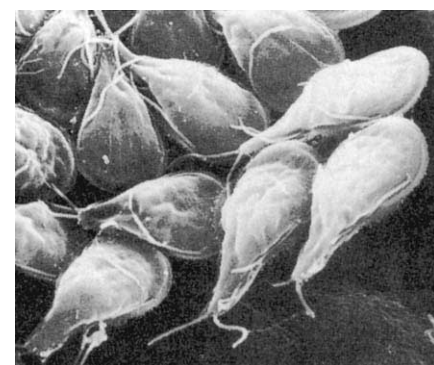


Fig. 2 Scanning electron micrograph of *Giardia* attached to the inner surface of the intestine, $\times 1,800$. (Courtesy of K. Vickerman.)

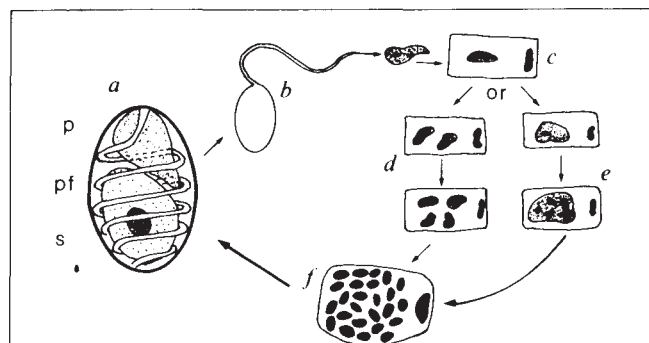


Fig. 1 The life cycle of a microsporidian such as *Nosema*, notorious for causing epidemic diseases in honeybees and silkworms. The feeding stage within the host cell is a minute amoeba, no bigger than some bacteria, with a nucleus but no mitochondria or peroxisomes, and which is structurally the simplest eukaryotic cell known. It multiplies inside the host cell by a simple mitotic mechanism (c-e). Eventually cell division ceases (f) and the cells differentiate into a unicellular resting spore (a) containing a coiled hollow tube, the polar filament (pf), a specialized secretory product of the Golgi apparatus which becomes obvious during spore formation. To infect a new host the polar filament turns inside out to form a hollow tube (b) like a flexible hypodermic needle, through which the infectious part of the spore (s) is extruded into its new host cell. The polaroplast (p), involved in polar filament extrusion, is absent from Rudimicrosporea. (From ref. 7, courtesy of M. Sleight.)

sitic habit, microsporidia themselves could not have been the host that engulfed purple bacteria to make the first mitochondrion. Phylogenetic analysis, based mainly on ultrastructural characters, suggests that this host was not an amoeboid cell like a microsporidian (and even less likely to be the wall-less prokaryote invoked by some speculators⁶), but was a swimming biciliated eukaryotic monad with a rigid cytoskeleton⁵. Its closest living relatives^{2,5} are the archezoan metamonad protozoa of the class Anaxostylea, whose best-known representative is the diplomonad *Giardia*, an important cause of dysentery in humans (Fig. 2).

Diplomonads include free-living species as well as gut symbionts, so are ecologically as well as cytologically good

bionts (gram-positive bacteria) that were taken up by the very same anaxostylean protozoan cell that first acquired mitochondria and chloroplasts. According to this view the biochemical symbiosis between mitochondria, peroxisomes and chloroplasts now apparent in plant cells was first established during the simultaneous concerted conversion of three different prokaryotic symbionts into new organelles, in an already highly evolved eukaryotic anaxostylean host, following a major bout of indigestion.

The fossil record suggests that these events took place only about 850 million years ago, and that the origin of the eukaryotic cell was between 1,500 and 850 million years ago. Vossbrinck *et al.* guess⁴ a much earlier date (about 2,800 million

years ago) for the divergence of microsporidia from other eukaryotes, but if this were so we would expect to see unequivocal evidence for undoubted eukaryotes very much earlier than we do. As the eukaryotic cell fundamentally depends on oxygen there is no good reason to connect the apparently later origin of mitochondria with the onset of an aerobic atmosphere. It is more probable that it was triggered, perhaps as much as 2,000 million years after oxygen became widespread, by the origin of phagocytosis and the first eukaryotic cell², before which cellular endosymbioses could not have occurred. Several archezoa use oxygen despite having no mitochondria; the anaerobic character of others is more likely to be a secondary adaptation to competition by the aerobically more efficient mitozoa than a relic from early eukaryotic evolution.

Work in progress in M. Sogin's laboratory will soon tell us whether *Giardia*, like microsporidia, diverged from mitozoa at a very early date and therefore whether the Metamonada are also primitively without mitochondria. Similar studies will show whether the Parabasalia primitively lack mitochondria, or whether, as I have suggested, their hydrogenosomes evolved from former mitochondria when their ancestor first colonized animal guts. Sequence data are even more important for the Archamoebae, which may represent either the most primitive eukaryotes, amoebae that have secondarily lost mitochondria or, very likely, a mixture. □

1. Cavalier-Smith, T. in *Endocytobiology II* (eds Schwemmler, W. & Schenk, H.E.A.) 265-279 (de Gruyter, Berlin, 1983).
2. Cavalier-Smith, T. *Ann. N.Y. Acad. Sci.* (in the press).
3. Vossbrinck, C.R. & Woese, C.R. *Nature* **320**, 287-288 (1986).
4. Vossbrinck, C.R., Maddox, J.V., Friedman, S., Debrunner-Vossbrinck, B.A. & Woese, C.R. *Nature* **326**, 411-414 (1987).
5. Cavalier-Smith, T. in *Progress in Phycological Research* (eds Round, F.E. & Chapman, D.J.) 309-347 (Biopress, Bristol, 1986).
6. Margulis, L. *Symbiosis in Cell Evolution* (Freeman, San Francisco, 1981).
7. Sleight, M.A. *The Biology of Protozoa* (Arnold, London, 1973).

T. Cavalier-Smith is a Reader in the Department of Biophysics, Cell and Molecular Biology, King's College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

Immunology

Antibody-antigen flexibility

Robert Huber and William S. Bennett

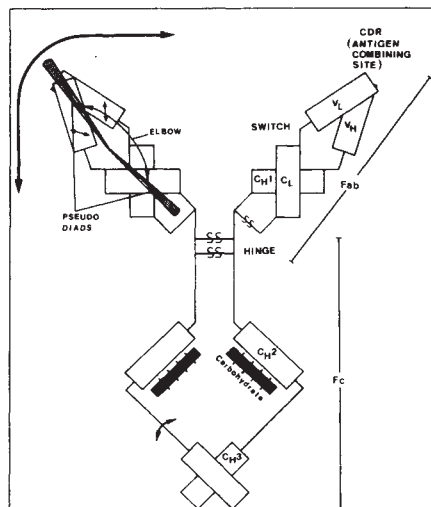
OUR view of the structural basis of antibody-antigen recognition and the role of flexibility in this process seems to evolve in oscillations whose amplitudes are asymptotically damped by the growing amount of reliable structural data. From Pauling's idea¹ of a very flexible and pluripotent antibody molecule, we have moved to a view of a molecule consisting of rigid regions of Fab fragments, light-chain dimers and the complexes of both with haptens (see ref. 2 for a recent review). The elucidation³ last year of the structure of a complex between lysozyme and a Fab fragment against lysozyme seemed to confirm the idea that both the antibody and the antigen retain their native structures on formation of the complex. Now, however, in a study reported on page 358 of this issue⁴, Colman *et al.* show limited structural adaptation of both components of another antibody-antigen complex, influenza virus neuraminidase and the NC41 Fab fragment against this viral antigen.

To understand the significance of this discovery, one should recall that, quite in contrast to the seemingly rigid dimeric modules formed by the globular antibody domains, the antibody molecule as a whole (consisting of four or more such modules linked like beads on a string) has served for years as a paradigm of molecular flexibility⁵. Segmental flexibility in antibodies has been demonstrated by various methods⁶⁻¹⁰, and later crystallographic studies defined in more detail the potential modes of motion by analysing the same or closely related molecules in different environments (see ref. 5 for a review).

The figure summarizes the various flexing modes of an immunoglobulin G molecule described to date. Angular motion of the Fab arms relative to each other and to the stem of the molecule, called Fc, is extensive, covering a range from maximal extension to almost parallel alignment. The arm movement is allowed by an extended segment of about 18 amino-acid residues linking the C_H1 and C_H2 domains, which is conformationally flexible but interrupted by a regularly folded segment rich in proline and cysteine residues. Similarly, elbow angles from 115 to 180° have been observed. Until now the V_H-V_L, C_L-C_H1, C_H3-C_H2 modules and the C_H2 domain (see figure) seemed always to move as rigid bodies, although recent analyses^{11,12} of light-chain dimers indicate a potential for variability in the V_L-V_L association. This potential now seems to be realized in the NC41 Fab-neuramin-

idase complex described by Colman *et al.*, which shows an uncommon V_H-V_L pairing.

Flexibility between the modules is functionally significant as it enables the bivalent antibody molecule to bind to epitopes (antigenic determinants) in various relative arrangements. It may also be involved



Generalized structure of an immunoglobulin G molecule, containing one Fc (stem) and two Fab (antigen-binding) fragments. The intact molecule has two heavy (H) and two light (L) chains. The constant (C) region of the H chains has three domains, C_H1, C_H2 and C_H3. V_H, variable region of the H chain; C_L, constant region of the L chain; V_L, variable region of the L chain; CDR, complementarity-determining segment. On each Fab fragment, V_L and V_H domains associate to form a variable module and C_H1 and C_L form the constant module. Double-headed arrows indicate the extent of potential motion of different parts of the molecule.

in the effector functions of antibodies, if sites of interaction with receptors or other macromolecular components in the Fc region are stabilized or exposed by Fab motion. The mode of V_H-V_L association, however, has a direct influence on the spatial arrangement of the complementarity-determining segments of heavy and light chains and thus determines binding specificity.

If such variability in the V_H-V_L arrangement of a given antibody indeed exists, it would add another dimension to antibody diversity by allowing a given antibody molecule to recognize a spectrum of epitopes. This potential might explain how antibodies raised against peptides derived from a protein can recognize the intact protein or, as pointed out by Colman *et al.*⁴, why some antibodies seem to cross-

react with completely unrelated antigens. A cautionary word is in order, however, as in none of the examples of uncommon V_H-V_L or V_L-V_L pairing is there experimental evidence that the same molecule is capable of forming the normal (or any other) variable-domain pairing. One must then wonder if the unusual pairings observed in these molecules are not simply their preferred structures. (After all, our prejudice as to what 'normal' variable-domain pairing constitutes is based on a sample of only five or six structures.) In the case of NC41 Fab, the amino-acid residues that were believed to mediate the V_H-V_L interaction are conserved. Thus, the interpretation of Colman *et al.* that the observed V_H-V_L pairing results from antigen binding seems plausible, as there is no obvious reason why the free antibody fragment should not exhibit the normal pairing.

The elbow angle of NC41 Fab is about 150°, indicating an open conformation and the absence of longitudinal non-bonded interactions between V_H and C_H1 (for a discussion of this issue, see ref. 13). It is similar in this respect to the lysozyme antibody-antigen complex³ and to the Kol immunoglobulin G which has properties of an autoantibody. The question of whether the elbow angle is influenced by antigen binding remains unsolved as the structure of free NC41 Fab is not known. Note, however, that the proposed V_H-V_L rearrangement could result in a change in elbow angle and in different longitudinal contacts, which might serve to propagate a signal from the variable to the constant module.

Although the evidence for a structural change in NC41 Fab on complex formation is indirect, the structure of a neuraminidase species closely related to the antigen in the complex described by Colman *et al.* is known in the free state, so that a more direct comparison is possible. The epitope recognized by NC41 Fab consists of four spatially adjacent but non-contiguous polypeptide segments of the neuraminidase molecule. It represents an exposed extended surface area not unlike the situation in the lysozyme antibody-antigen complex³. The polypeptide segments of neuraminidase in contact with the Fab are slightly deformed by the

1. Pauling, L. J. *Am. Chem. Soc.* **62**, 2642-2657 (1940).
2. Davies, D. R. & Metzger, H. A. *Rev. Immun.* **1**, 87-117 (1983).
3. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Science* **233**, 747-753 (1986).
4. Colman, P. M. *et al. Nature* **326**, 358-363 (1987).
5. Bennett, W. S. & Huber, R. *CRC Crit. Rev. Biochem.* **15**, 291-384 (1984).
6. Noelken, M. E., Nelson, C. A., Buckley, C. E. & Tanford, C. J. *biol. Chem.* **240**, 218-224 (1965).
7. Yguerabide, J., Epstein, H. F. & Stryer, L. *J. molec. Biol.* **51**, 573-590 (1970).
8. Shumaker, V. N. *et al. Molec. Immun.* **17**, 413-423 (1980).
9. Pecht, I. in *The Antigens Vol VI* (ed. Sela, M.) 1-62 (Academic, New York, 1982).
10. Cathou, R. E. *Compr. Immun.* **5**, 37-83 (1978).
11. Furey, W., Wang, B. C., Yoo, C. S. & Sax, M. J. *molec. Biol.* **167**, 661-692 (1983).
12. Chang, C. H. *et al. Biochemistry* **24**, 4890-4897 (1985).
13. Huber, R. *Science* **233**, 702-703 (1986).

interaction, an observation that is not surprising for exposed loops. Interestingly, a residue that is known to be involved in the catalytic action of neuraminidase alters its position drastically on complex formation. This structural change could explain the inhibition of neuraminidase activity by NC41 Fab, which does not directly block the catalytic site. Crystallographic work

on other antigen-antibody complexes is under way, which may help to answer some of the questions raised by Colman *et al.* □

Robert Huber is at the Max-Planck Institut für Biochemie, 8033 Martinsried and William S. Bennett is at the Max-Planck Institut für Molekulare Genetik, Ihnestr. 63-73, 1000 Berlin 33, FRG.

Organic ferromagnets

Polymers from the Soviet Union

Richard Friend

THE possibility of constructing a ferromagnetic array of spins on radicals in an organic material has been known for some time^{1,2}. Small polyradical molecules have been found that do show strong ferromagnetic exchange coupling between the electron spins within the molecule³. It has not until now been possible either to arrange ferromagnetic coupling between such molecules or to extend them, for example as polymer chains. However, recent work by Ovchinnikov and colleagues in the Soviet Union, reported on page 370 of this issue⁴, shows that a polymer based on polydiacetylene, with nitroxyl biradical side groups, behaves as a ferromagnet, with a Curie temperature of about 150 °C.

The exchange interaction between two spatially localized magnetic spins is usually antiferromagnetic. There are only a handful of examples of non-metallic ferromagnetic salts, in contrast to the thousands of known antiferromagnetic transition metal salts. The reason for this is that the antiferromagnetic spin singlet combination of two interacting spins can

occupy a single spatial quantum state, whereas the Pauli principle requires that the ferromagnetic spin triplet occupies higher energy levels. The group of materials where there is scope for building structures with these interactions reversed are organic molecules with extended π electron systems. Many of these molecules have been studied in the past decade for their ability to delocalize charge and behave as metals or high mobility semiconductors. Thus, conjugated polymers such as polyacetylene can show metallic levels of conductivity when doped, as I have recently discussed in a News and Views article⁵, and among the many metallic molecular charge transfer salts there are even some which are superconductors⁶.

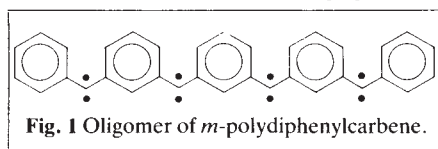


Fig. 1 Oligomer of *m*-polydiphenylcarbene.

occupy a single spatial quantum state, whereas the Pauli principle requires that the ferromagnetic spin triplet occupies higher energy levels. The group of materials where there is scope for building structures with these interactions reversed are organic molecules with extended π electron systems. Many of these molecules have been studied in the past decade for their ability to delocalize charge and behave as metals or high mobility semiconductors. Thus, conjugated polymers such as polyacetylene can show metallic levels of conductivity when doped, as I have recently discussed in a News and Views article⁵, and among the many metallic molecular charge transfer salts there are even some which are superconductors⁶.

The same properties that encourage charge delocalization (extended π mol-

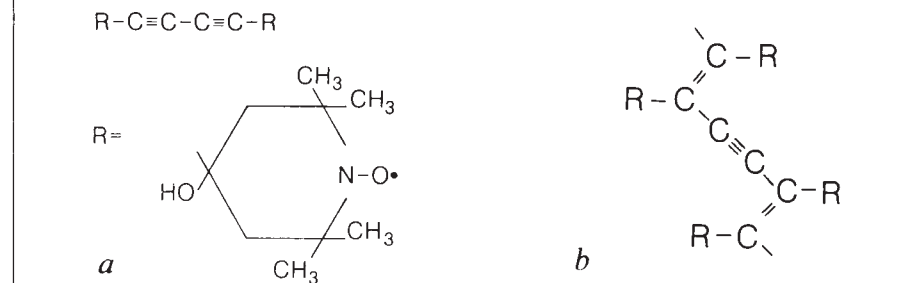


Fig. 2 a, The diacetylene monomer and stable radical side groups (R) used by Ovchinnikov's group^{4,8}. b, Polydiacetylene backbone.

ecular orbitals along a polymer chain, many molecular wavefunctions with similar energies) are also desirable for building materials with interesting magnetic properties. J. Torrance⁷ has developed a straightforward scheme to explain the magnetic interactions between spins in these systems and has shown how it is possible to arrange them to be ferromagnetic. The overlap of wavefunctions on adjacent sites can be considered to mix into the ground-state wavefunction the excited state obtained by transferring the electron spin from one site to the other. If the lowest excited level is non-degenerate it must necessarily be the antiferromagnetic singlet. But if it is degenerate, Hund's rules give the triplet to be lower in energy and ferromagnetic ordering is preferred. These conditions should be met for polymers with π electrons along the chains which are odd-alternate, and very detailed experimental work by Iwamura and co-workers³ has been done on model oligomeric systems of this type. The oligomer of *m*-polydiphenylcarbene shown in Fig. 1, for example, has a total of 8 spins (p_z and non-bonding electron at each bridge site), which are ferromagnetically coupled to give a ninefold degenerate (nonet) ground state. Unfortunately, the larger oligomers have not yet been made, and the coupling between adjacent oligomer molecules in the solid is antiferromagnetic. The way forward is to extend the chains as polymers. Torrance has seen ferromagnetic behaviour in the reacted product of 1,3,5-triaminobenzene and

iodine a few times, and he conjectures that this compound is a partially oxidized *meta*-linked phenyl polymer.

Ovchinnikov and co-workers^{4,8} have followed a different route, and have attached stable radicals as side groups to the diacetylene monomer shown in Fig. 2a. They then polymerized the monomer with heat or exposure to ultraviolet radiation to give a polydiacetylene backbone, shown in Fig. 2b. Whereas the spins on the radicals on the monomer show Curie-like behaviour, those present in the polymer are ferromagnetically coupled. Ovchinnikov and colleagues consider that the exchange interaction in the polymer is mediated by the π electrons along the polymer chain. At present only a small fraction (0.1 per cent) of the monomer is converted to ferromagnetic polymer, although this fraction can be magnetically separated from the residue and shows a respectably

high Curie temperature of 150 °C or above.

These new results are far from ideal, with a poor reaction yield for the synthesis and a poorly characterized product, but they will certainly stimulate further efforts to synthesize other organic ferromagnets. The discovery of ferromagnetism now completes the list for organic materials of those electronic properties (metallic conduction, superconductivity and ferromagnetism) formerly associated only with inorganic materials. It will be some time to come, however, before organic ferromagnets find applications in place of traditional magnetic materials. □

1. McConnell, H.M. *J. chem. Phys.* **39**, 1910 (1963).
2. Ovchinnikov, A.A. *Theoret. Chim. Acta*, Berlin **47**, 297 (1978).
3. Sugawara, T., Bandow, S., Kimura, K., Iwamura, H. & Itoh, K. *J. Am. Chem. Soc.* **108**, 368 (1986).
4. Korshak, Y.U., Medvedeva, T.V., Ovchinnikov, A.A. & Spector, V.N. *Nature* **326**, 370 (1987).
5. Friend, R.H. *Nature* **322**, 308 (1986).
6. *Proc. Int. Conf. Synthetic Metals, Kyoto in Synthetic Metals* **17**, (1986).
7. Torrance, J. in *Proc. Int. Conf. Synthetic Metals, Kyoto in Synthetic Metals* **17**, (1986).
8. Korshak, Y.U., Ovchinnikov, A.A., Shapiro, A.M., Medvedeva, T.V. & Spector, V.N. *Pis'ma Zh. Eksp. Teor. Fiz.* **43**, 309 (1986) (English edn *JETP Lett.* **43**, 399; 1986).

Richard Friend is at the Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK.

Developmental neurobiology

Trophic factor theory matures

J.W. Lichtman and P.H. Taghert

THE Nobel prize in Physiology and Medicine was awarded last September to Stanley Cohen and Rita Levi-Montalcini, recognizing the importance of their work, begun in the 1950s, on the identification and purification of nerve growth factor (NGF), a protein essential for the survival of several neuronal cell types in the peripheral nervous system. By extrapolation, NGF is thought by many neurobiologists to be a representative example of a hypothetical class of agents that mediate an obligatory trophic dependence of neurons for the targets they innervate. As first suggested by Levi-Montalcini and Viktor Hamburger, neuronal survival may depend on successful competition for target-derived trophic factor that is in limited supply. Over the past few years the trophic hypothesis has been strengthened by the demonstration that the targets of NGF-sensitive neurons do indeed synthesize this protein^{1,2} and by the equally important finding that NGF-sensitive neurons have specific receptors for NGF³⁻⁵. The details of the trophic dependence, however, including the exact role and mechanism of NGF action, remain obscure. On page 353 of this issue⁶, Davies *et al.* contribute a noteworthy study that focuses on the details of trophic factor and trophic factor receptor expression. Davies *et al.* address fundamental questions regarding the 'when' and the 'where' of NGF action, and their answers are in many ways surprising.

Differentiation

The differentiation of neurons occurs in phases of growth (proliferation, axon outgrowth and synapse formation) and regression (cell death and synapse elimination). Before the work of Davies *et al.* the potential scope of NGF action ranged throughout nearly the entire life of neurons. With the exception of neuronal proliferation, which seems to be independent of it, NGF has been implicated in virtually all other phases of neuronal differentiation. For example, NGF puffed out of a pipette⁷ or injected intracerebrally *in vivo*⁸ can redirect the orientation of growing axons towards the source of the agent. Hence, during axon outgrowth it was not unreasonable to think that NGF may be a chemotactic agent that guides growth cones up a concentration gradient towards their targets. Also at issue is the site of NGF action: non-target tissues such as Schwann cells appear to be NGF-immunoreactive^{9,10}. Identifying the site of expression is clearly important to the understanding of when and where

neurons first come into contact with NGF.

Davies *et al.*⁶ examined NGF and NGF-receptor expression in early embryos of mice at a stage when ingrowing sensory fibres of the trigeminal ganglion first reach the maxillary process and developing whisker pad. Asking when NGF messenger RNA and NGF could first be detected, they find, surprisingly, that the onset of transcription and subsequent translation does not precede but is actually coincident with the arrival of the first sensory axons. That detectable levels of the messenger RNA or NGF itself are not present before the arrival of the first axons strongly suggests that NGF action begins only after target innervation; that messenger RNA levels are also temporally correlated allows for the unexpected possibility that axons elicit the synthesis of NGF in their targets. These data provide strong evidence that NGF has little if any role in guiding growth cones to their appropriate targets.

But what of neuronal receptivity to NGF? Equally important in defining the time period of NGF action is to determine when neurons first express NGF receptors on their surfaces. Using an ¹²⁵I-labelled NGF probe, Davies *et al.* report that the receptors are not detectable on sensory neurons until after they arrive at their target. This finding is consistent with the observation in tissue culture that developing neurons are, for a while, not dependent on NGF. The picture is emerging that NGF begins to mediate survival after neurons have contacted their prospective targets and not before.

This interpretation critically depends on the sensitivity of the assays used to measure NGF and NGF-receptor expression. In the case of specific RNA and protein assays, the levels of sensitivity meet reasonable expectations: for example, the two-site immunoassay can detect less than a picogram of NGF (500 times more sensitive than earlier procedures)¹¹. In the case of the assay for receptor levels, performed with labelled NGF and autoradiography, there is reason to suspend judgement as to the exact time of receptor expression. An immunoprecipitation assay¹² using a monoclonal antibody against NGF receptor turned out to be 50 times more sensitive than detection using directly labelled NGF.

Two final points are worth mentioning. First, Davies *et al.* demonstrate NGF messenger RNA expression in the developing whisker pad by *in situ* hybridization techniques. This simple finding establishes, by the most direct means to date, that the

target is at least one source of NGF. It also counters suggestions that trophic support derives exclusively from Schwann cells⁹: Davies *et al.* detect significant concentrations of the messenger RNA within the cutaneous epithelium where there are no Schwann cells. Second, trigeminal ganglion explants from periods of development that precede axon outgrowth nevertheless show NGF-receptor expression along growing axons at the appropriate time. Hence, despite lack of contact with the target, neurons are capable of entering the NGF-dependent phase.

Time of action

The work of Davies *et al.* merits attention because it places NGF expression and receptivity within a coherent cellular and developmental context. The results show that the time of NGF's action ensues only after target innervation. This further suggests that it is in the realm of competition between neurons and of long-term survival that NGF action must be explained. Finally, the temporal details provided suggest a more complex interpretation of neuronal development than previously thought. Before these studies, a broader view of the scope of NGF action from the period of neurogenesis to that of competition prevailed: now it appears that neurons initially have intrinsic capabilities that enable them to reach the stages of competitive interactions, or perhaps that they require yet other factors from elements of the cellular environment in which they proliferate and through which their axons grow.

It is curious why nature chose to use a protein factor — NGF — to promote the survival of a sub-population of neurons. How is NGF action mediated; what are the biochemical and physiological effects of NGF on individual neurons; what are the consequences of those effects on long-term survival; and does this sort of interaction hold for all neurons? The simple yet elegant measurements of Davies *et al.* help to refine the role of NGF, to define where and when to look for other trophic factors and help to elucidate what these factors are doing. □

1. Heumann, R. *et al.* *EMBO J.* **3**, 3183-3189 (1984).
2. Shelton, D.L. & Reichardt, L.F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7951-7955 (1984).
3. Banerjee, S.P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **70**, 2519-2523 (1973).
4. Frazier, W.A. *et al.* *J. biol. Chem.* **249**, 5513-5519 (1974).
5. Chandler, C.E. *et al.* *J. biol. Chem.* **259**, 6882-6889 (1984).
6. Davies, A.M. *et al.* *Nature* **326**, 353-358 (1987).
7. Gundersen, R.W. & Barrett, J.N. *J. Cell Biol.* **87**, 546-554 (1980).
8. Menesini-Chen, M.G. *et al.* *Archs Ital. Biol.* **116**, 53-84 (1978).
9. Rush, R.A. *Nature* **312**, 364-367 (1984).
10. Finn, P.J. *et al.* *J. Neurocytol.* **15**, 169-176 (1986).
11. Korsching, S. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3513-3516 (1983).
12. Taniuchi, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **83**, 1950-1954 (1986).

J.W. Lichtman and P.H. Taghert are in the Department of Anatomy and Neurobiology, Washington University School of Medicine, St Louis, Missouri 63110, USA.

Nomenclature of immunological markers

SIR—Your editorial policy¹ concerning the nomenclature of markers of human leukocyte differentiation antigens is welcome, but does not go far enough. A more extensive list has been official nomenclature of the International Union of Immunological Societies and the World Health Organisation since 1983, and is now widely accepted. This nomenclature emerged in the First Workshop on Human Leucocyte Differentiation Antigens², held in Paris in 1982, and was extended by the Second Workshop³ (Boston, 1984). The report of the Third International Workshop held in Oxford in October 1986 is now being printed and a summary of the updated recommended nomenclature is available in ref. 4; lists of antibodies in each cluster will be found in refs 2 and 3.

It is not clear why you have decided to stop at CD8 rather than adopt the more extensive and growing list. And why retain interleukin-2 (why not TAC?) receptor rather than CD25? We do not discourage the use of the former, but encourage the use of the latter in conjunction with trivial names, including the name of the specific antibody in use. The reasons are more obvious in CD25 than in other cases. CD25 recognizes the putative β -chain (55K) of the receptor, and there is now strong evidence to implicate a second α -chain (75K)⁵. Use of the same chain by different molecular complexes is no longer a bizarre observation, and the workshop nomenclature attempts to avoid the attendant complications.

We were somewhat disappointed by other aspects of your policy. It has given undue weight to the reagents of two major commercial companies, rather than perhaps better ones from other companies and laboratories, or a better scientific judgement. For instance, three sets of CD1 (CD1a, CD1b and CD1c) are now recognized and both antibodies you chose as examples are representatives of CD1a. Furthermore, none of them was the first to be produced nor represents the best available. More seriously, you give TL as the mouse CD1 equivalent, when you recently published evidence that this is unlikely to be correct⁶.

C. MILSTEIN

For the International Workshops
on Human Leucocyte Differentiation

MRC Laboratory of Molecular Biology,
Hills Road, Cambridge CB2 2QH, UK

1. *Nature* 325, 660 (1987).

2. *Leucocyte Typing* (eds Bernard, A. et al.) (Springer, Heidelberg, 1984).

3. *Leucocyte Typing II* (eds Reinherz, E. L., Haynes, B. F., Nadler, L. M. & Bernstein, I. D.) (Springer, Heidelberg, 1986).

4. Shaw, S. *Immunol. Today* 8, 1–3 (1987).

5. Smith K. A. *Immunol. Today* 8, 11–13 (1987).

6. Calabi, F. & Milstein, C. *Nature* 323, 540–543 (1986).

Soliton theory and Jupiter's great red spot

SIR—Antipov *et al.*¹ recently presented some remarkable results from an ingenious laboratory analogue of the shallow-water equations on a rapidly-rotating β -plane, in which an apparently isolated, stable anticyclonic eddy formed spontaneously in the presence of a mechanically driven lateral shear flow. The latter was claimed to be a 'physical analogue' of the jovian great red spot with certain advantages over other competing models, such as the one presented by us² a few years ago which was based on sloping thermal convection in a differentially-heated, rotating fluid. In comparing their work with our experiments, however, Antipov *et al.* wrongly concluded that our results would apply only to deep fluids ($D/L > 1$, where D is the fluid depth and L a horizontal length scale), and could exhibit no dispersion in eddy propagation.

Although our experiments were carried out in a container of aspect ratio greater than 1, the dynamically significant measures of that aspect are with respect both to the so-called 'Prandtl ratio of scales'³, f/N (where f is the Coriolis parameter and N the Brunt-Väisälä or buoyancy frequency⁴), and to a condition relating to the applicability of hydrostatic equilibrium. Thus, $D/L \leq f/N$ for the occurrence of baroclinic eddies (coincidentally placing them in a similar regime to the shallow-water eddies of ref. 1 with respect to the relevant deformation radius). A similarly scaled aspect ratio in the jovian atmosphere would result in $D/L \ll 1$, owing merely to the respective values of f and N . For the applicability of hydrostatic equilibrium, a scale analysis of the vertical momentum and thermodynamic equations (A. A. White, personal communication) indicates that D/L should be $\ll Ri^{1/2}$ (where the Richardson number $Ri = N^2 D^2 / U^2$, and U is a horizontal velocity scale). As Ri is about 10^3 in our experiments, the aspect ratio used ($D/L \approx 2.5$) is sufficiently shallow for the fluid to remain in hydrostatic equilibrium (the consequences of departures from hydrostatic equilibrium in baroclinic instability were considered by Stone⁵).

The absence of dispersion in rotating annulus experiments, suggested by Antipov *et al.*, refers only to idealized flows in systems of constant depth. Residual potential vorticity gradients arising from mean zonal flow curvature, for example, will result in practice in the ubiquitous presence of weak dispersive effects⁶, while the use of sloping endwall boundaries⁷ can allow a wide variation in externally imposed 'planetary' vorticity gradients analogous to the effects of spherical curvature in a planetary atmosphere. The effects of such sloping boundaries are currently

being investigated in the system studied in ref. 2, and preliminary results indicate only small quantitative changes in the propagation speeds of eddies and the location of flow regimes in parameter space from our earlier work.

The application of the analogue of Antipov *et al.* to the jovian atmosphere rests principally upon the formal mathematical similarity between the shallow-water equations and those describing barotropic (or equivalent barotropic) motion in a continuous atmosphere⁸. The extension of their work to a 'baroclinic' atmosphere, discussed by Nezlin⁸, using eigenfunctions of the vertical structure equation with non-zero eigenvalues, can only be regarded as partial because, if only one such mode were present, the eddy would still be incapable of transferring heat, and hence of releasing potential energy stored in the mean zonal flow (the essence of baroclinic eddy generation⁹), unless a continuous vertical phase tilt (associated with a complex vertical structure function) were present. The relative importance of barotropic and baroclinic effects on Jupiter is currently uncertain as, contrary to the apparent claims of Antipov *et al.*¹, there is little direct evidence of the mean zonal flows 'feeding' the great red spot on Jupiter and maintaining it against the inevitable effects of friction⁹.

In comparing their results with other quasibarotropic models of the great red spot, Antipov *et al.* cite the soliton model of Maxworthy and Redekopp¹⁰, which is fundamentally a weakly nonlinear asymptotic reduction of the problem to a solution of the well-known Korteweg-de Vries equation. In contrast, the observation¹ that the flow speeds in their laboratory eddies exceed the phase speed c indicates a strongly nonlinear phenomenon, probably more akin to the 'intermediate geostrophic' eddies in the shallow-water numerical models of Williams and Yamagata^{11,12}. Like the eddies of ref. 1, the latter also exhibit an asymmetry between cyclonic and anticyclonic eddies, arising from nonlinear divergence associated with the displacement of the free upper surface. As remarked by Williams^{11,12}, the applicability of the nonlinearly divergent 'intermediate geostrophic' model to a continuously-stratified baroclinic fluid remains controversial and, in any event, also raises questions as to the existence of the cyclonic jovian 'barges' and persistent cyclonic features at other latitudes⁹. The main point of similarity between the results of Antipov *et al.* and our own work seem to be in the way horizontal advective nonlinearities can evidently conspire to steepen eddies which arise from an instability of a mean zonal flow, resulting in compact configurations which are robustly stable against competing effects such as dispersion and

dissipation.

Further laboratory and theoretical studies of such features are likely to yield much valuable insight into the general properties of nonlinear waves and eddies, whether or not the nature of the jovian great red spot ultimately becomes apparent.

P.L. READ

*Geophysical Fluid Dynamics Laboratory,
Meteorological Office (21),
Bracknell, Berkshire RG12 2SZ, UK*

1. Antipov, S.V., Nezhlin, M.V., Snezhkin, E.N. & Trubnikov, A.S. *Nature* **323**, 238-240 (1986).
2. Read, P.L. & Hide, R. *Nature* **302**, 126-129 (1983); **308**, 45-48 (1984).
3. Prandtl, L. *Essentials of Fluid Dynamics* (Blackie, Glasgow, 1952).
4. Gill, A.E. *Atmospheric-Ocean Dynamics* (Academic, New York 1982).
5. Stone, P.H. *J. Fluid Mech.* **45**, 659-671 (1971).
6. Hide, R. & Mason, P.J. *Geophys. Astrophys. Fluid Dyn.* **10**, 121-156 (1978).
7. Mason, P.J. *Phil. Trans. R. Soc. A278*, 397-445 (1975).
8. Nezhlin, M.V. *Soviet Astr. Lett.* **10**, 221-226 (1984); *JETP Lett.* **34**, 77-80 (1981).
9. Read, P.L. *Icarus* **65**, 304-334 (1986).
10. Maxworthy, T. & Redekopp, L.G. *Icarus* **29**, 261-271 (1976).
11. Williams, G.P. & Yamagata, T. *J. Atmos. Sci.* **41**, 453-478 (1984).
12. Williams, G.P. *J. Atmos. Sci.* **42**, 1237-1243 (1985).

Glycine-binding sites and NMDA receptors in brain

SIR—I was fascinated by the report by Johnson and Ascher¹ in which they describe an interaction between the amino-acid neurotransmitters glycine and glutamate in cultured mouse neurons. Glycine is generally accepted as an inhibitory transmitter, certainly within the spinal cord, whereas glutamate is an excitatory transmitter. The effect of glycine, they report, was not blocked by strychnine, which antagonizes the Cl⁻ mediated inhibitory responses to glycine.

Their data provide a rationale for an observation I made after my colleagues and I published on the presence and distribution of strychnine-insensitive ³H-glycine binding sites in supraspinal regions of the rat brain observed using autoradiography². Comparison of our data with the findings of Monaghan and Cotman³ on the binding of ³H-glutamate to N-methyl-D-aspartate (NMDA) sites showed that the distributions of glycine and NMDA binding were identical. This is readily demonstrated by comparing Fig. 1 from ref. 2 with Fig. 3 from ref. 3. The autoradiograms are very similar, perhaps most strikingly so in the hippocampus, where the greatest density of both NMDA and glycine sites is present in the CA1 radiatum and oriens layers. The only apparent discrepancy in the hippocampus occurs in the stratum lucidum which was labelled by ³H-glutamate but not by ³H-glycine. However, these glutamate binding sites are kainate-sensitive and not NMDA sites³. The similarity between glycine and NMDA extends to the distribution of sites labelled

by the non-competitive NMDA antagonist ³H-MK801 (ref. 4).

Thus the effect of glycine, as suggested by Johnson and Ascher, may be mediated through receptors located externally on neuronal membranes and associated with the NMDA receptor. The autoradiographic data certainly support the view that the NMDA receptor complex includes modulatory sites analogous to those in the GABA_A receptor complex. Interestingly, the modulatory sites previously described in the NMDA receptor complex seem to have endogenous substrates within the brain, namely Mg²⁺ (ref. 5) and alpha-endopsychosin⁶. Presumably we can now add glycine to this list.

N. G. BOWERY

*Neuroscience Research Centre,
Merck Sharp and Dohme,
Terlings Park, Harlow,
Essex CM20 2QR, UK*

1. Johnson, J. W. & Ascher, P. *Nature* **325**, 529-531 (1987).
2. Bristow, D. R., Bowery, N. G. & Woodruff, G. N. *Eur. J. Pharmac.* **126**, 303-308 (1986).
3. Monaghan, D. T. & Cotman, C. W. *J. Neurosci.* **5**, 2909-2919 (1985).
4. Bowery, N. G. & Hudson, A. L. *Br. J. Pharmac.* **89**, 775P (1986).
5. Nowak, L., Bregestovski, P., Ascher, P., Herbert, A. & Prochiantz, A. *Nature* **307**, 462-465 (1984).
6. Quirion, R. *et al. Peptides* **15**, 967-973 (1984).

Can a negative quantity be deemed a probability?

SIR—One could say that it is reasonable to use the idea of negative probabilities¹⁻⁴ if a body of results valid for normal (that is, positive) probabilities is changed in an elegant or informative way when some of the probabilities involved are given negative values. It is for this sort of reason that one introduced negative energy states, negative temperatures and made similar extensions of concepts previously regarded as constrained to have positive numerical values. Hence, although the frequency definition of probability seems to rule out negative probabilities, there could nevertheless be rules obeyed by probabilities in physics or in probability theory itself, which could make it elegant or convenient to use the extended definition. Suggestions for this approach have been made and have in the main been based on quantum theory. The purpose of this letter is to add to the recent discussion¹⁻⁴ by giving a class of simple examples based on thermodynamics.

I have shown elsewhere⁵ that the first and second law of thermodynamics applied to two identical bodies of constant heat capacity, *C*, lead one to the well-known inequality for the means

$$\frac{1}{2} (T_1 + T_2) \geq (T_1 T_2)^{1/2} \quad (1)$$

where *T*₁, *T*₂ are the temperatures of the two bodies prior to equilibration. Equilibrium by contact yields the left-hand side of

(1) as the final temperature. Equilibration by running a reversible Carnot engine between the bodies yields the right-hand side as the final temperature. Thus we have, reasonably, that constant-entropy equilibration yields a lower temperature than constant energy equilibration. If the bodies have unequal heat capacities (*C*₁, *C*₂ say), then the same procedure leads to the inequality for the weighted means

$$pT_1 + qT_2 \geq T_1^p T_2^q \quad (p + q = 1) \quad (2)$$

where *p* and *q* are formally probabilities, whose interpretation is brought about by the physical argument based on the first and second law of thermodynamics:

$$p = \frac{C_1}{C_1 + C_2}, \quad q = \frac{C_2}{C_1 + C_2} \quad (3)$$

The result (2) is correct mathematically provided the condition *p*, *q* > 0 is added. The positivity and normalization makes *p* and *q* normal probabilities.

Next, apply the first and second law to two bodies of different constant heat capacities *C*₂ < 0 < *C*₁, noting that negative heat capacities occur for black holes, stars and other gravitationally bound systems. We could for simplicity consider the case of small temperature changes so that the assumption of constant (temperature-independent) heat capacities is reasonable. One now finds

$$pT_1 + qT_2 \leq T_1^p T_2^q \quad (p + q = 1) \quad (4)$$

If *C*₁ + *C*₂ < 0 we have *p* < 0 < *q*, while *q* < 0 < *p* if *C*₁ + *C*₂ > 0. If *T*₁ = 2, *T*₂ = 3 then an example for the former case is provided by *C*₁ = 1, *C*₂ = -2 leading to 4 < 9/2, and for the latter case by *C*₁ = 2, *C*₂ = -1 leading to 1 < 4/3.

Just as in (2) one can proceed by abstraction from a physical result to a mathematical result, so one can proceed similarly with respect to (4). One arrives at the 'reversed' inequality (4) for means⁶ which are now weighted by one positive and one negative 'probability'. The introduction of negative probabilities is of course not essential even in this case, but at least it does not seem to be too artificial, because of the simple analogy between (2) and (4). We shall explore this matter in a more general setting elsewhere⁷.

P.T. LANDSBERG

*Faculty of Mathematical Studies,
University of Southampton,
Southampton SO9 5NH, UK*

1. Mückenheim, W. *Phys. Rep.* **133**, 337 (1986).
2. Maddox, J. *Nature* **320**, 481 (1986).
3. Parry, D.E. *Nature* **321**, 644 (1986).
4. Mückenheim, W. *Nature* **324**, 307 (1986).
5. Landsberg, P.T. *Phys. Lett.* **67A** 1 (1978); **78A**, 29 (1980); *J. Math. Anal. Appl.* **76**, 209 (1980).
6. Pečarić, J. *Glas. Mat. (Zagreb) Ser. III* **16**, 229 (1981); **19**, 235 (1984).
7. Landsberg, P.T. & Pečarić, J. *Phys. Rev.* (in the press).

Presentation of the past

Alan Costall

The Shaping of Modern Psychology. By L.S. Hearnshaw. Routledge & Kegan Paul: 1987. Pp. 408. £19.95, \$49.95.

It is odd to find a psychologist, however courageous, attempting to write a general history of the subject, especially a work that spans "the dawn of civilization to the present day". Not only is there now so much — too much, perhaps — to be included, but the task itself can hardly be approached with the innocence of earlier days. The historians, motivated not merely by professional jealousy but real misgivings about our past efforts, insist that history is too serious a matter to be entrusted to anyone but themselves.

Their charges against the scientist-historian are various. One is 'internalism', the persisting treatment of science as though (as Engels put it) it had fallen from the skies, a self-contained activity quite isolated from any wider social context. Another is 'Whiggism', where the past is portrayed as a hesitant yet inevitable progress towards its culmination in our times and our ways. A further charge (there are others) seems to be most dominant in recent years: 'presentism'. Although this is often taken as a synonym of Whiggism, it is, as I understand it, distinct and perhaps defensible. In this case, the past is drawn upon selectively to provide a context for the discussion of current issues. On a limited scale, this is surely what all scientists do in the literature review preceding their research reports. One might question whether presentism, so construed, counts as history at all; but, if adopted explicitly, I cannot see that it must lead to *bad* history.

Professor Hearnshaw is a psychologist who many years ago, and before most others, specialized in the history of his subject. His book, *A Short History of British Psychology, 1840–1940* (Methuen, 1964), was well received by historians and psychologists alike, and his recent biography of Cyril Burt (Hodder & Stoughton, 1979) is regarded as the official verdict on Burt's unfortunate genius for inventing IQ data and even research associates. Despite the opportunities, however, there are few hints of scandal in his new book. Even judged on its own terms as a "synopsis" of the develop-

ment of modern psychology, the scope is impressive, not only in time (in the order of tens of thousands of years) but also diversity. For example, there are authoritative treatments of theories of mind in Eastern religions and of Soviet psychology, and an insightful evaluation of Freud's contribution. The author also



Classification in mind — a thirteenth-century attempt to relate Aristotle's division of the rational soul to the cerebral ventricles.

dwells, to good effect, on the tensions within psychology which arise from its links with both the natural and the moral sciences, including history itself.

Hearnshaw's approach, however, is more questionable. He shows little respect for the recent controversies and criticism within the history of science, and despite the extensive use of primary sources there is surprisingly little reference to the historical work of his fellow psychologists or to their current concerns. He does not even seem to feel any obligation to appease the historians for embarking upon yet another presentist and largely internalist history. In fact, what is most striking is the way in which Hearnshaw shamelessly turns the historians' criticism of scientists' history on its head. As he explains at the outset, the central purpose of his history is to lay to rest the spectre, raised by the historians, that psychology

might, after all, be just one darned thing after another. His argument is blunt: psychology is progressive because a serious Whig history can indeed be written:

A study of the history of psychology refutes the contention that the subject has not progressed. ... If today psychologists shiver in a 'winter of discontent' ... it must be because they have forgotten their history, have lost touch with the inspiration it provides, and are out of tune with the slow march of time [pp. 5–6].

Not surprisingly, there is much talk of psychology's "true task" and "ascent", and plenty of pioneers and father-figures: Plato, the father of cognitive psychology; Augustine, the first, great introspective psychologist; Hobbes, the ancestor of "one strand" of modern psychology. However, although Professor Hearnshaw takes pride in psychology's eventual arrival as a lively scientific discipline, he is neither complacent nor unaware of its lack of cohesion. He is critical, for example, of the confusion and one-sidedness of the 'cognitive revolution' and its resort to the computer metaphor:

A machine cannot think, any more than a book can remember. The meaning of the words on a printed page are bestowed by minds; they have only a delegated intentionality. ... Computer bewitchment threatens to eclipse an appreciation of the biological and historical depths of human nature [p. 272].

Professor Hearnshaw is not, therefore, engaged in a glorification of the *status quo*. But his is a Whig history nevertheless, for he portrays the past as an extended preparation for what he sees as our future task: the unification of psychology.

This is a scholarly and thoughtful book. Yet, as a new history, there is something strikingly unmodern about its approach, a lack of a sense of the profound reflexivity of the human sciences. Certainly, there is a recognition of the importance of history in psychology, an appreciation that many apparent universal laws of human nature may, in fact, be specific to particular social conditions. But Hearnshaw neglects the crucial twist. Psychology is now itself a part of our culture, and hence an influence upon the way we conduct our affairs. Hearnshaw's use of history is fine for its avowed aim, to exhort and rally the troops. But its recurrent appeal to an abstract and reified 'past' leads to just the kind of history which obscures rather than clarifies the human nature of our science. □

Alan Costall is a Lecturer in the Department of Psychology, University of Southampton, Southampton SO9 5NH, UK.

Science of helium in technology

P.V.E. McClintock

Helium Cryogenics. By Steven W. Van Sciver. Plenum: 1986. Pp.429. \$69.50, £55.60.

LIQUID helium is something of an oddity. Its existence as a liquid at all is rather marginal, as shown by the ease with which it can be vaporized by tiny influxes of heat — just one watt is enough to evaporate about a litre of liquid in an hour. For temperatures below 2.17K, it behaves as though it were an interpenetrating mixture of two completely miscible fluids: a (relatively ordinary) normal fluid component, and a superfluid component which carries no entropy and whose viscosity is identically zero. It is the latter component that gives rise to liquid helium's celebrated frictionless-flow properties, enabling it, for example, to climb out of any open vessel in which it is placed.

After spending the first several decades after its discovery in relative obscurity, liquid helium has quite suddenly emerged as an industrial commodity of considerable technological importance. The question now arises as to how this strange,

ephemeral and rather fragile liquid can most efficiently and safely be employed by engineers as an industrial coolant, particularly for large superconducting magnets.

Professor Van Sciver's *Helium Cryogenics* sets out to provide comprehensive answers to this question. The book is aimed at graduate students and professionals in cryogenic engineering, low-temperature physics and superconductivity who need to make use of liquid helium in the course of their work. The intention is to provide an up-to-date description of the physics of the substance, and then to see how this knowledge can be applied to a variety of engineering applications. After some helpful scene-setting introductory material, Chapters 3 and 4 provide the necessary physical descriptions. It is the following three chapters, treating heat transfer to static and flowing fluid helium in its superfluid, normal, gaseous and mixed-phase states that form the vital core of the book. Chapter 8 discusses helium liquefaction/refrigeration, and the final chapter deals rather briefly with a medley of special topics including helium dilution refrigeration, and cooling by adiabatic demagnetization.

The particular strengths of the book, to my mind, lie in its unified coverage of what has become a very large field, in its excellent central accounts of heat transfer and transport, and in the valuable graphs

and tables of numerical data. There are one or two minor disappointments that should also be mentioned. First, some of the historical comments in the earlier chapters are off-key; as one example, the discovery of the superfluidity of liquid helium is not normally attributed to the Leiden group. Secondly, Chapter 4, on helium as a quantum fluid, is less than up-to-date and contains some unfortunate lacunae. For instance, the otherwise excellent summary of Landau's excitation model makes no mention of Landau's famous explanation of superfluidity, or of the Landau critical velocity. And although a full and detailed treatment is given of the phenomenological 2-fluid model, its close connection to the Landau excitation model is neither stated nor demonstrated. It seems particularly regrettable that the section on Bose condensation omits any mention of the verification by Sears *et al.* (*Phys. Rev. Lett.* **49**, 279; 1982) of the existence of a Bose condensate in liquid helium.

Taken as a whole, however, *Helium Cryogenics* succeeds admirably in its central aims. It should be of real assistance to any scientist or engineer wishing to make cryogenic use of helium down to about 1K. □

P.V.E. McClintock is a Reader in the Department of Physics and a member of the Low-Temperature Group, University of Lancaster, Lancaster LA1 4YB, UK.

JUST PUBLISHED OF URFS AND ORFS

A Primer on How to Analyze Derived Amino Acid Sequences

by

Russell F. Doolittle, University
of California, San Diego

7" x 10", 112 pages, paperback only,
\$14.00 U.S.

Special Offer for Prepaid Orders

All prepaid orders receive a 10% discount off list price. Please add \$1.85 for postage and handling. California residents add 6% sales tax. Orders to:

Marketing Director
UNIVERSITY SCIENCE BOOKS
20 EDGEHILL ROAD
MILL VALLEY, CA 94941

Tracking back to Africa

D.R. Higgs

The Sickled Cell: From Myths to Molecules. By Stuart J. Edelstein. Harvard University Press: 1986. Pp.197. \$25, £22.25.

SICKLE cell disease is a common and serious genetic disorder, affecting approximately one in every five hundred newborn black children. Even with good medical care this disease carries a relatively high mortality, but in rural Africa most affected individuals die in infancy. Those patients who do survive have a chronic anaemia and may suffer painful sequelae as a consequence of blockage of the microvascular system by sickled red blood cells.

Since the discovery by Linus Pauling and Vernon Ingram that it results from a single amino-acid substitution in the β -haemoglobin (Hb) chain, sickle cell anaemia has become the prime example of a molecular disease. In 1971 President Nixon proclaimed that "something should be done about it", and by 1973 Congress had authorized a multi-million dollar pro-

ject to tackle the problem. Although there was an element of scientific hubris in all this, such an approach had worked before; the successful Apollo programme had started with a similar Presidential declaration. Unfortunately, it has not worked out so well with sickle cell disease. Over the past 20 years we have discovered the precise underlying molecular defect, and we now understand in exquisite detail the mechanism by which molecules of HbS polymerize and damage the red blood cells, but this knowledge has had little impact on our ability to treat afflicted patients.

Stuart Edelstein has made many contributions to our understanding of the sickling process, most importantly by elucidating the molecular structure of the sickle polymer. I suspect that like most research workers in this field he is frustrated by the discrepancy between our knowledge and our ability to apply it to the treatment of patients. In writing *The Sickled Cell*, he made a bold attempt to break the mould of the past 20 years by returning to the 'scene of the crime', Africa, to evaluate the myths and folklore surrounding sickle cell disease in the hope that therein lies the molecular pharmacologist's dream. Karl Popper, who argued that all — or very nearly all — scientific theories originate from myths,

and that a myth may contain important anticipations of scientific theories, would doubtless approve of this approach. Indeed, as Edelstein argues, many invaluable drugs, such as aspirin, morphine and digitalis, have evolved from mythical potions to become carefully prescribed therapeutic agents.

I came to this book anticipating some innovative and long-overdue lateral thinking about sickle cell disease, but rapidly became disillusioned. The first half of the book deals with the emergence of the sickle mutation in its broader anthropological context, and includes sections on primate evolution, population genetics and the history of Africa. The second half describes the molecular basis of sickling. This is well-trodden territory and I felt both aspects were dealt with rather superficially. The myths surrounding sickle cell disease are alluded to throughout the book, but again I was disappointed by the lack of fresh material; arguments that cassava, yam and fava beans may have a modifying effect on the disease are neither new nor compelling. The less well known story of 'repeater children', a widespread myth throughout Africa, is examined in detail. According to various African traditions these children return repeatedly to the same parents. They die prematurely only to arrive again with the birth of the next child. Although this could be the animistic interpretation of recurrent infant mortality due to a genetic disease, there is such a high background of childhood illness of all types in this environment that any connection between these *ogbanje* children and sickle cell disease seems, at best, tenuous.

From the book, it appears that African mythology hardly abounds with stories based on sickle cell disease. Perhaps Edelstein's quest should have been extended to similarly affected cultures in Arabia, India and the Mediterranean. Unfortunately, from my own superficial search of their fables, the most relevant of them tells of the revenge taken by Mother Earth on the all-powerful Greek god Uranus — he was castrated with a sickle.

The concept behind this book is good, but it seems to me that it was written prematurely, with insufficient data and drawing upon ideas that have not yet been fully worked out. On the other hand, it provides a sobering lesson for those in pursuit of the cause of other important genetic diseases. Even when the genes have been found and the proteins characterized, as in sickle cell disease, it is not necessarily the end of the story. These other researchers, too, may have to resort to desperate tactics to find a cure for their patients. □

D. R. Higgs is Clinical Research Fellow in the Medical Research Council Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK.

On the heart beat

Andrew P. Somlyo

Cardiac Muscle: The Regulation of Excitation and Contraction. Edited by R.D. Nathan. Academic: 1986. Pp.324. \$69.95, £58.50.

Cardiac Muscle. By E.D. Canale, G.R. Campbell, J.J. Smolich and J.H. Campbell. Springer-Verlag: 1986. Pp.318. DM460.

OF THE inner organs of beasts and fowls, none has been relished more, by arts and science alike, than the heart. Reputed of old as the seat of emotions, but better known these days for its vulnerability to the ravages of food and smoke, the interest of modern granting agencies in this simple pump has contributed much to our understanding of cell physiology and to a dramatic decline in deaths from cardiac disease. These two volumes make useful further contributions to our understanding of 'affairs of the heart'.

Nathan's book is a multi-author mosaic of 12 chapters by as many authors and co-authors that reflects the *Zeitgeist* of modern cardiac physiology. The emphasis, in general, is on excitation, rather than contraction; electrophysiology and electrophysiological methods are dominant, and only a slim chapter, dealing with calcium-binding proteins, gets the reader much past the surface membrane and sarcoplasmic reticulum. Within its range, however, the book is an authoritative survey by some of the protagonists in the field.

Sodium-calcium exchange is the main theme of three chapters, dealing with its electrogenicity and contribution to contractile regulation. Fortunately, controversy is not fully avoided, as attested elsewhere in the book in the discussions of the role of calcium bound to the surface membrane. On balance, I am inclined to believe, with the majority, that while such bound calcium may play a significant role in the regulation of surface membrane properties, it is unlikely to be mobilized at sufficiently high concentrations to make much of a contribution to the direct activation of contraction.

Wandering off the beaten path, the discussion of the electrophysiological effects of peptide hormones provides an interesting vista, moving with caerulein to depolarize the pancreas and nerve growth factor prodding the Na-K pump. The chapter dealing with the release of calcium from the sarcoplasmic reticulum, by an author who played the dominant role in developing the concept of calcium-induced calcium release as the trigger of excitation-contraction coupling in the heart, could serve as a model of objectivity. It also concludes that the spectacular evolution-

ary progress of the inositol 1,4,5-trisphosphate bandwagon has, alas, been stopped short by forces of nature from being the direct mechanism of excitation-contraction coupling in the heart or, as other evidence suggests, in skeletal muscle.

The book is an educational and entertaining 'browse', but I feel obliged to mention the plaintive "Note" of a contributor that "Due to the lengthy process of publication, much new information... is not included"; this accompanies a chapter containing a subheading that proclaims of acetylcholine-modulated K-channels: "Intracellular Messenger Unlikely". The role of G proteins has been recognized since this volume appeared, and the laboratory of the author concerned has demonstrated that similar mechanisms mediate muscarinic effects on calcium currents. To lovers of the printed word, such examples of delayed information transfer (some manuscripts submitted in 1984 but published only in 1986), herald the inevitability of electronic publication in fast-moving fields.

The other volume reviewed here aims "to provide a comprehensive and up-to-date review of the structure of the heart", and succeeds admirably. The excellent electron micrographs remind one that the quality of illustrations, rightly or wrongly, greatly influences our judgement of structural studies. The book also has a

BOOKS ON NUCLEAR ENERGY

for the specialist
for the public

For our free catalog
call or write:

AMERICAN
NUCLEAR
SOCIETY

555 N. Kensington Avenue
La Grange Park, Illinois 60525
312-352-6611

cohesiveness usually attained only by a single author or, as in this case, by close collaborators who obviously devoted considerable effort to its planning and execution.

The morphology of cardiac muscle, its specialized conduction system and innervation are described in depth worthy of a reference text, yet without being boringly encyclopaedic. In addition to the more classical aspects of light and electron microscopy of cardiac muscle, the volume is particularly strong in its discussion of the development of the myocardium and, as befits a field in which the authors have considerable expertise, the properties of cultured cardiac muscle cells. In a book containing 101 illustrations and over 500 references, those interested in the fashionable can learn about the atrial spe-

cific granules storing the atrial natriuretic peptides, while the discussion of corbular sarcoplasmic reticulum and of the avian cardiac conduction system will please the bird lovers.

Reminded of a comment attributed to one of the Marx Brothers — "Whom are you going to believe, me or your eyes?" — most of us will be inclined to trust the visual evidence. Books such as this one will strongly reinforce our faith in our eyes, and in the importance of learning about structure to understand function. It is highly recommended, whether for reading or reference, and should remain useful for years to come. □

Andrew P. Somlyo is Director of the Pennsylvania Muscle Institute, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

The dimensions of art

J.D. Mollon

The Forms of Color: The Interaction of Visual Elements. By Karl Gerstner. MIT Press: 1986. Pp. 179. \$29.95, £29.95.

THIS is a flirt of a book. Lured along by many handsome illustrations, the reader is led down instructive by-ways in the history of geometry and ornament; but the intellectual thread grows ever more slender, the book turns into a vehicle for displaying the author's own paintings, and there is no final consummation to be had.

The questions that Gerstner sets himself are interesting ones. Can we have a system of form, derived from primitives, in the way we can have a well-ordered space that represents all possible colours? How does colour affect the perceptual organization of form? And are there systematic correspondences — intra-modal synaesthesias — between particular colours and particular forms?

After taking Wyszecki's rhombohedral lattice as an example of a modern colour-order system, Gerstner passes on to Diquat's "Méthode pour faire une infinité de desseins différens" of 1722, and thence, via the history of group theory and fractal geometry, to the geometric art of Islam. A mosaic from the Hall of the Ambassadors in the Alhambra is used to show how a simple algorithm will generate a pattern rich in rhythms and alternative percepts (see the accompanying illustration).

The same theme is continued in a discussion of Wilhelm Ostwald's "Harmonie der Formen". Although Ostwald is well known for his colour system, his system of form is largely forgotten, and Gerstner gives a detailed summary. Ostwald starts

with those regular polygons that can fill a plane so as to leave no space over when juxtaposed: the equilateral triangle, the square and the hexagon. Suppose that we conceptually subdivide one of these basic forms into equal parts and we connect with real lines a chosen subset of the intersections that have been generated by the subdivision. If we now duplicate this unit form many times by translation or mirror reversal, there will emerge a complex superordinate pattern that it is difficult to imagine in advance. Ostwald's algorithms are extended in the paintings of the Swiss artist Hans Hinterreiter: here the unit forms are duplicated on to topologically distorted grids and so the range of emergent patterns is multiplied.

When Gerstner finally comes to ask

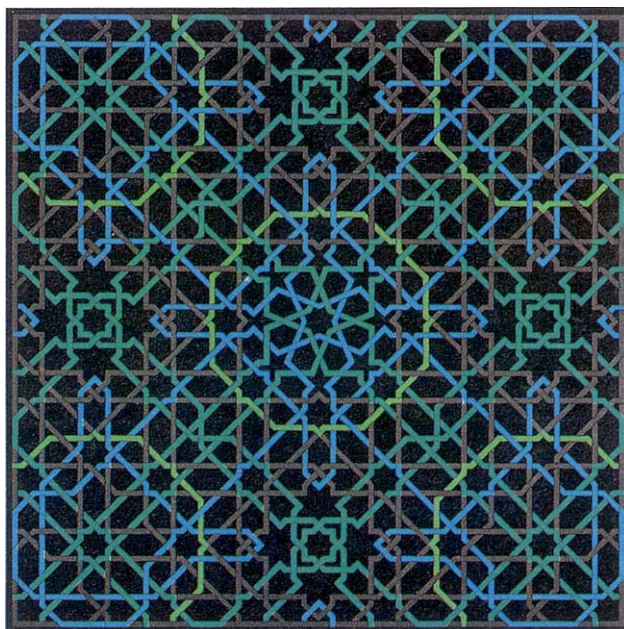
whether there are primitives of form, he abandons geometry and selects forms that he judges to be elementary in their emotional expression. An eight-pointed star, like the colour yellow, is active, warm and out-radiating; a circle, like the colour blue, is passive, cold and in-radiating; a square standing on its corner, like the colour red, is tense and loud.

Although Karl Gerstner is an unusually learned artist and although his book, if it is about anything, is about human perception, very little is actually said about the experimental psychology of form, about empirical studies of synaesthesias or about the physiological properties of our visual system. There is nothing about modern studies of perceptual organization; or about the degree to which form and colour are independently analysed within the visual system; or about texture, although texture as a surface property of objects is closely analogous to colour, and all textures could probably be represented in a space of dimensionality not vastly greater than that of colour space.

But perhaps it is as well that Gerstner does not add to his intellectual baggage the primal sketch of David Marr and the textons of Bela Julesz. From the pointillists to Hockney there is a tradition of artists giving theoretical commentaries on their work. But almost always the rationalizations are opaque or mistaken. The art must stand or fall by itself. And Gerstner's agreeable compositions are best judged for their own sake. □

J.D. Mollon is a Lecturer in the Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

Pattern of a mosaic in the Alhambra. By colouring the interleaved strips of the original, Gerstner draws attention to some of the many competing organizations that the eye can impose upon the pattern. (Reproduced from The Forms of Color.)



The sombre view of AIDS

Malcolm Rees

Extrapolation from short-term data is risky, but the development of AIDS among those infected with HIV by blood transfusions, if properly described by a normal distribution, suggests a 15-year mean incubation and a high incidence of disease among carriers of the virus.

THE nature of the AIDS (acquired immune deficiency syndrome) epidemic is of considerable interest to us all, but so far the underlying epidemiology of the disease is not understood. Consequently, predictions of the incidence of the disease can be reliable only for perhaps one to three years ahead. Since the identification of the AIDS-related virus, or human immunodeficiency virus (HIV), and the introduction of tests for the antibodies thereto, interest has begun to focus on the relationship between infection and the onset of AIDS itself.

Until recently, it was thought that only a small proportion of those infected, perhaps only 10 per cent, would move on to full AIDS, while another group would suffer from a variety of non-fatal illnesses including persistent generalized lymphadenopathy (PGL) and the more serious AIDS-related complex (ARC).

This view has now been criticized by a number of writers who have proposed that, in any year, those infected have a certain probability of contracting AIDS, and that this process will continue into the future. For example, Holmes writes¹ that the "annual risk of AIDS among seropositive individuals approximates 2 per cent to 4 per cent a year, and the risk of AIDS-related illnesses is considerably higher". Lawrence² mentions one estimate, described as "probably conservative", that "7 per cent of the currently infected but still healthy individuals will develop AIDS each year". Blattner³ states that "over several years of follow-up, the risk of AIDS developing in seropositive homosexual men ranged from 8 per cent to 10 per cent per year in New York to 4 per cent to 6 per cent in San Francisco".

It would seem that these estimates cannot describe the generality of HIV infections in the United States because if there are, as has been claimed, between 1 million and 2 million seropositive individuals in the United States, we would expect 100,000 new cases of AIDS each year, roughly seven times as many as the present estimate of the number of new cases each year by the Centers for Disease Control (CDC). Even if we assume that the majority of the 1 to 2 million seropositives have been infected only recently,

so that there may not yet have been time enough for them to have contracted AIDS, these very high percentages cannot be reconciled with the actual numbers of AIDS cases. They are almost certainly derived from samples biased towards patients showing early signs of disease.

These writers imply that AIDS cases are being generated from a cohort of infected persons according to some probability distribution. The assumption that, from a given cohort of seropositives, a fixed number will develop AIDS each year, implies a rectangular distribution. If a constant proportion of the survivors in the cohort develops the disease each year, on the other hand, a geometric distribution would be appropriate, suggesting that new AIDS cases in a cohort of infected people decline each year, as more of those infected go down with the disease. We know from the San Francisco data that the reverse of this is the case.

Frequency

The purpose of what follows is to investigate other distributions describing the probability that a person infected at a specific time will develop AIDS at later times. In particular, I shall consider the familiar normal or gaussian distribution, which is uniquely specified by its mean μ (the average time between infection and the emergence of the disease) and the standard deviation σ . There are some theoretical reasons why the log-normal distribution may be more appropriate⁴, but it is not possible to distinguish between these two distributions on the basis of the few years of data which are now available.

That the normal or log-normal distribution may describe the frequency of AIDS cases in time after infection is suggested by our knowledge of other diseases. Christiansen *et al.*⁵, in a study of the time of onset of diabetic nephropathy after the diagnosis of early-onset diabetes, show a log-normal pattern with a mean interval between the two diagnoses of more than 20 years. Parry⁶, studying scrapie in sheep, gives a cumulative distribution of cases in time after birth which appears normal, though with some positive skewness.

Once the underlying distribution is known, the numbers of cases to be expected in future years arising from those already infected may be calculated from a knowledge of the numbers infected at various times in the past. Comprehensive estimates of the future incidence of the disease would also, of course, require estimates of the numbers of new cases of infection in the future.

People infected in a particular year may be regarded as a cohort that will generate AIDS cases in subsequent years, so that the number of new cases arising in a future year from that cohort alone will be a product of the size of the cohort and the probability of developing AIDS, given by the probability distribution. Because some of those infected may escape AIDS, the integral of this distribution over time will be equal to, or less than, unity, the difference indicating the proportion of those infected who will not contract the disease. For the whole of a community, the number of cases arising in a particular year will be the sum of the calculated contributions of HIV cohorts of each year.

Table 1 Transfusion-associated AIDS cases by years of infection and diagnosis

Year of infection	Year of diagnosis of AIDS*							Totals
	1978/9	1979/80	1980/1	1981/2	1982/3	1983/4	1984/5	
1978	0	0	0	0	0	2	1	3
1979		0	0	2	1	5.5	4.5	13
1980			0	0	8	11.5	5	24.5
1981				1	7.5	8.5	17	34
1982					3	14.5	20.5	38
1983						4.5	20.5	25
1984							7	7
Totals	0	0	0	3	19.5	46.5	75.5	144.5

*Total excludes 1985. Source: ref. 7.

Table 2 Hypothetical numbers of AIDS cases according to time between infection and illness

Year of infection	Year of illness							Totals
	1978/9	1979/80	1980/1	1981/2	1982/3	1983/4	1984/5	
1978	0.07	0.12	0.20	0.32	0.50	0.73	1.06	3.00
1979		0.45	0.79	1.33	2.15	3.33	4.96	13.01
1980			1.38	2.41	4.05	6.53	10.12	24.49
1981				3.27	5.71	9.58	15.45	34.01
1982					6.70	11.69	19.61	38.00
1983						9.11	15.89	25.00
1984							7.00	7.00
Total	0.07	0.57	2.37	7.33	19.11	40.97	74.09	144.51

Assumes normal distribution with $\mu=15$; $\sigma=5$.**Table 3** Actual cases of transfusion-related AIDS compared with 'theoretical numbers'

Year		1978/9	1979/80	1980/1	1981/2	1982/3	1983/4	1984/5	Total absolute differences
Actual number		0	0	0	3	19.5	46.5	75.5	
Theoretical numbers									
$\mu=6$	$\sigma=2$	0.02	0.20	0.99	3.87	12.87	36.83	89.72	32.60
$\mu=9$	$\sigma=3$	0.03	0.28	1.35	4.91	15.07	38.79	84.14	24.35
$\mu=12$	$\sigma=4$	0.05	0.41	1.84	6.12	17.22	40.08	78.79	17.41
$\mu=15$	$\sigma=5$	0.07	0.57	2.37	7.33	19.11	40.97	74.09	14.67
$\mu=18$	$\sigma=6$	0.09	0.73	2.86	8.33	20.51	41.52	70.44	20.06
$\mu=21$	$\sigma=7$	0.10	0.84	3.28	9.23	21.65	41.67	67.72	25.21

Unfortunately, we do not know the numbers infected in each year, but only the total number of those who develop AIDS. But if the probability distribution were known, it would be possible to estimate the numbers infected in each year by solving for x in $A \cdot x = y$ where A is a matrix whose elements are the probabilities that members of a particular cohort will develop AIDS at 1 to n years into the future, x is a vector of the number of infections by year and y is a vector of AIDS cases by year of diagnosis.

The problem may be simplified if the A matrix and the x vector are modified so that x is replaced by z , a vector of the number of HIV infections in any one year who will eventually get AIDS and A by the matrix B whose elements are the probabilities that people with HIV leading to AIDS will develop the disease in future years.

Each column of this matrix must sum to unity because, by hypothesis, by the year n , all those infected will develop AIDS. This adjustment of the equations is tantamount to excluding from consideration all HIV infections that will not lead to AIDS.

The task now becomes that of solving for z in $B \cdot z = y$; our ignorance of the proportion of infections that will not lead to AIDS has been by-passed.

AIDS time distribution

Although there are some studies, such as the San Francisco cohort study, which are concerned with HIV infection in a sample as a function of time, it is by no means easy to estimate actual times of infection from cohort information. Some of the cohort will almost certainly be infected at recruitment and, without intensive and continu-

ous monitoring, it will not be possible to tell when new infections in fact occur.

In the San Francisco case, members of the cohort were originally selected in 1978, 1979 and 1980 for an investigation of hepatitis B infection, and were followed up in 1984. The study does not make it possible to pinpoint the time of infection accurately because some of the subjects were already infected when first approached and because the time interval between the first and second approaches was between 4 and 6 years. Seropositive rate estimates for 1981 to 1983 were interpolated from the figures for 1980, and cannot be regarded as reliable.

However, there is a sample of AIDS cases where the time of infection is known accurately. Peterman *et al.* have reported⁷ cases of people who have contracted HIV infection from blood transfusions, and who have subsequently developed AIDS, when the exact time of infection may be determined by testing stored blood. These samples do not, of course, include those who may have been infected by the same route who have not yet developed AIDS or who may never do so.

AIDS cases associated with transfusions are shown in Table 1, by year of

infection and of diagnosis. Data for 1985, incomplete in the report by Peterman *et al.* are excluded, for which reason the total number of cases is reduced from 155 to 144 (cases straddling the years have been assigned half in each).

To decide which normal distribution most closely fits these data, normal distributions with various parameters were fitted to the figures in Table 1. I conclude (see below) that the best estimate of the mean μ is 15 years, and that the standard deviation σ amounts to approximately 5 years. Because of the comparatively small number of transfusion cases these parameters may need amendment in future.

The implications of the large estimated value of the mean time between infection and overt disease are considerable. Thus we know (from the data in Table 1) that by 1984/85 there had been 34 cases of AIDS among those known to have been infected by transfusion in the United States in 1981. But the normal distribution with a mean and standard deviation of 15 years and 5 years respectively would by that year have produced only 1.257 per cent of all the AIDS cases that will ultimately be attributable to 1981 infection by the transfusion route. This in turn implies that there must have been a total of 2,705 AIDS infections by transfusion in the United States in 1981. Summing over all the years of infection, the same data suggest that there will eventually be 23,155 cases of transfusion-linked AIDS, assuming that no more people are infected by this route (and that those infected do not die of some other disease first).

With the same parameters, it is also possible to calculate for each year of infection the numbers of AIDS cases to be expected among those infected by transfusion; the results, which should be compared with the data in Table 1, are shown in Table 2. Table 3, on the other hand, gives the numbers of expected cases with different assumptions about the values of the mean and the standard deviation defining the normal distribution. It will be seen that the best fit, as measured by the absolute difference scores, is given by the values $\mu = 15$ years and $\sigma = 5$ years. In reality, there is a slightly better fit with $\mu = 14$ years, but in the calculations below I have used the value of 15 years both because the difference between that and 14

Table 4 Matrix of probabilities of AIDS onset by year for cases infected in different years

Year of AIDS onset	Year of infection						
	1978	1979	1980	1981	1982	1983	1984
1978/9	0.00121	0	0	0	0	0	0
1979/80	0.00211	0.00121	0	0	0	0	0
1980/1	0.00354	0.00211	0.00121	0	0	0	0
1981/2	0.00571	0.00354	0.00211	0.00121	0	0	0
1982/3	0.00885	0.00571	0.00354	0.00211	0.00121	0	0
1983/4	0.01318	0.00885	0.00571	0.00354	0.00211	0.00121	0
1984/5	0.01887	0.01318	0.00885	0.00571	0.00354	0.00211	0.00121

Assuming a normal distribution with $\mu=15$, $\sigma=5$ for the distribution of AIDS cases through time.

Table 5 AIDS cases in the United States and the United Kingdom

United States	
Year	Cases
1978/79	5
1979/80	27
1980/81	128
1981/82	538
1982/83	1,827
1983/84	3,957
1984/85	7,358
United Kingdom	
Year	Cases
1981/82	1
1982/83	13
1983/84	40
1984/85	142
1985/86	269

years is small when account is taken of the sparsity of the data and of the circumstantial evidence that AIDS following transfusion-linked infection may develop rather earlier than among other infected groups, perhaps because of the relative frequency of children (known to succumb sooner) among those infected by transfusion.

Incidence of infection

On the basis of a normal distribution with $\mu = 15$ years and $\sigma = 5$ years, it is now possible to solve the matrix equation to derive the number of infected cases from recorded numbers of overt cases of the disease. The starting-point for the calculation is the matrix **B**, whose elements are probabilities derived from a normal distribution with the specified parameters. Very early infections have been ignored in the calculations as numerically unimportant.

United States. The **y** vector (the number of AIDS cases per year, mid-year to mid-year) is given in Table 5 (where the data are numbers of reported cases, mid-year to mid year). Table 6 shows the number of new infections estimated to have arisen by the end of each annual period and the last column shows the total number of infections (derived by summing the values in the previous column).

United Kingdom. The same procedure can be repeated for the United Kingdom, where data are recorded according to the time of notification rather than time of first diagnosis (Tables 5,6). Since the middle of 1986, AIDS cases have been recorded by the three-monthly period in which patients have first sought medical advice. The data in Table 5 refer to annual periods ending at the end of July each year. (The three notifications in 1982 have been divided one to 1981/82 and two to 1982/83).

Discussion

This model for the distribution in time of the occurrence of AIDS after infection

appears to give sensible results when generating estimates of the numbers of new infections in the United States and the United Kingdom. My estimate of the prevalence of infection in the United States based on the use of the normal distribution embodied in matrix **B** and published data for the incidence of new AIDS cases, leads to the estimate that at the end of 1984 or thereabouts, there were 2.5 million HIV infections that will result in AIDS over the next 30 years or so. This estimate is greater than that of Sivak and Wormser⁸, who estimated that there were 1.75 million people in the United States infected in mid-1985, but not drastically so.

The estimate of 100,000 AIDS infections in the United Kingdom in the middle of 1985 is very much greater than previous estimates. There seems to be a general opinion that there were about 30,000 infections in the middle of 1986, between a quarter and a third of my estimate for a year earlier. The discrepancy may be even greater than it appears because of the use of the AIDS notifications in the records for the United Kingdom rather than of

may be that, because the US AIDS statistics refer to the time of diagnosis (and are usually for this reason subject to upward revision retrospectively), the most recent figures will certainly underestimate the true state of affairs because of late notification and may also be depressed by mis-diagnosis or failure to notify because patients and physicians are seeking to avoid the stigma of AIDS.

In Britain, it seems that new infections may have fallen considerably between 1984 and 1985. An obvious implication of the model is that deaths from AIDS, both in the United States and in the United Kingdom, would continue to increase into the late 1990s even if the rate of new infection could be reduced to zero immediately. This assumes that there will be no treatment for those already infected. AIDS cases may rise more rapidly than deaths if survival time from AIDS to death increases in the future.

If the log-normal distribution proves to be more appropriate than that described here, it would be expected that the number of new cases of AIDS would begin to increase more sharply in the next few

Table 6 AIDS infections in the United States and the United Kingdom

United States			
Year	New infections	Total infections	
1978	4,132	4,132	
1979	15,108	19,240	
1980	67,350	86,590	
1981	263,482	350,072	
1982	751,897	1,101,969	
1983	714,901	1,816,870	
1984	669,596	2,486,466	
United Kingdom			
Year	New infections*	Total infections	
1981	826	826	
1982	9,306	10,132	
1983	14,412	24,544	
1984	61,098	85,642	
1985	23,646	109,288	

*Note that UK infections are derived from case notifications. This may cause small year-on-year anomalies in infection rates.

diagnoses as used in the United States; notifications may lag diagnoses by several months.

Unfortunately, the model carries the implication that HIV infection can usually be expected to lead to AIDS, but only over a period of many years.

The same calculations also reveal one hopeful element — the apparent decline in the growth of infection. Both in the United States and the United Kingdom, the absolute number of new infections per year would appear to have fallen in the most recent years for which the calculations are meaningful. Indeed, if the calculation had been applied literally to the data for the United States in 1985/86, the result would have been a slightly negative number of new infections in 1985, which is of course impossible. The explanation

years, and that the peak in the number of cases of AIDS would occur sooner than in the normal model.

I should like to thank Professor Stewart Cameron of Guy's Hospital Medical School and Dr John Seale for advice. □

Malcolm Rees is at the John Hunter Clinic, St Stephen's Hospital, Fulham Road, London SW10 9TH, UK.

1. Holmes, K.K. *J. Am. med. Ass.* **254**, 2254–2257 (1985).
2. Lawrence, J. *Scient. Am.* Dec. 1985: 70–79.
3. Blattner, W.A. *et al. Ann. intern. Med.* **103**, 665–670 (1985).
4. Aitchison, J. & Brown, J.A.C. *The Log Normal Distribution* (Cambridge University Press, 1957).
5. Christiansen, J.S. *et al. Diabetic Nephropathy*, **4**, 104–107 (1985).
6. Parry, H.B. *Scrapie Disease in Sheep* (Academic, New York, 1983).
7. Peterman, T.A. *et al. J. Am. med. Ass.* **254**, 2913–2917 (1985).
8. Sivak, S.L. & Wormser, G.P. *New Engl. J. Med.* **313**, 1352 (1985).

Science meets Islam in the Saudi desert

Oil revenues have bought Saudi Arabia many of the tools of science and technology. The challenge now is to build up a national science community that can apply them.

Riyadh

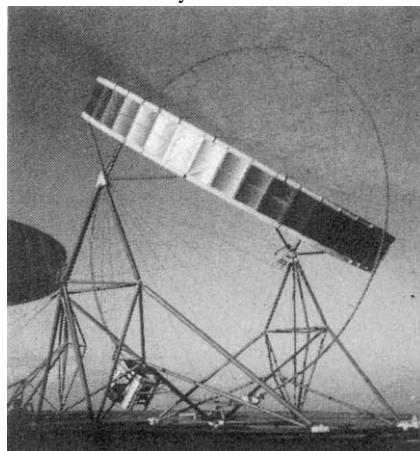
ONE of the most basic of statistics against which to measure the progress of any country is its population. But if the Saudis know what their population is, as surely they must, they are not saying. Semi-official suggestions range up to 11 million but the British embassy reckons 5–6 million Saudi Arabians and perhaps 2 million non-Saudis and yet other guesses are that there are no more than 2–3 million nationals. The reasons for the official reticence are not hard to find. First, a proper census would emphasize how small a population there is to defend such a large land mass should the Kingdom ever be threatened by, say, Iran. And, second, it would reveal just how dependent Saudi Arabia still is on external experts.

The latter problem, at least, is on the wane as the country's investments in higher education begin to pay off. Not least of these is the King Saud University which, all told, has an enrollment of 38,000 students. Of these, 24,000 (all male because the Islamic creed forbids co-education) are on a 2-year-old campus on the outskirts of Riyadh, where no expense seems to have been spared. Instruction in medicine, science and engineering is in English. At first this was because most of the teaching staff were from overseas, says Dr Salim Babhair, vice-dean of the dental school. But it is likely to stay that way because, to whatever extent 'Saudiization' of the staff proceeds, most of the teaching material will be in English.

So far, says Babhair, spending on education and health services has been protected from the effects of falling oil revenues that have at least slowed down progress in other sectors. Not only is all education free but there are still substantial allowances for university students and cash bonuses upon graduating. Surprisingly, not even primary school education is compulsory but it is reckoned that everybody completes a secondary education and that illiteracy exists only among the older, unschooled population. Quite what people read is a different matter, thanks to a censor whose black pen drew a veil over all but the head of Michelangelo's *David* when it appeared in a front-page photograph in the *International Herald Tribune* (7–8 March). Evolution is not taught at school and "is taught but only as a theory in which we do not believe" in the universities, according to an official of the King Abdulaziz City

for Science and Technology (KACST).

It is there that the state of development of scientific research in Saudi Arabia becomes somewhat apparent. The 'city' — known until two years ago as the Saudi Arabian National Center for Science and Technology, and still not much more than a small building housing more administrators than scientists — essentially controls the research programme of the country. It is administratively attached to the office of



Blue-skies research: testing solar devices is a major Saudi project.

the Prime Minister, King Fahd bin Abdul Aziz Al Saud, and has the task of funding and coordinating national and international research. In 1985–6 it awarded 29 grants worth SR 17.5 million (£3.2 million) and was supporting six national research projects at twice that cost. Although many of KACST's plans are in an embryonic state, collaborative work on solar energy is well advanced, even if the Saudi role is still largely to test devices made elsewhere. A joint US–Saudi programme, Soleras, has been running for ten years and has various projects including a solar-powered desalination plant (Riyadh is supplied mainly with desalinated water piped hundreds of miles from the coast) and Solar Village, 30 miles northwest of the capital. There, 160 arrays of 256 solar cells track the sun and routinely produce 350 kilowatts of direct current electricity during peak sunshine periods. The economics, however, are still such that the system would not be installed anywhere that the main grid can easily reach. On the same site are two experimental FRG–Saudi solar devices, each consisting of a parabolic collector focussing on a Stirling engine that drives a 50 kilowatt generator.

The cutback in oil revenues inevitably

poses some threat to the research budget. For example, the National Observatory Project calls for a major new optical telescope but, says KACST vice-president Dr Abdullah Al-Kadhi, it is no longer absolutely certain that the money will be available to build it. In any case, there is more difficulty than anticipated in finding a site that is not beset by problems of atmospheric dust; the best remaining candidate site will be examined later this month. One of the most important tasks of the National Observatory Project — and an example that is given of science and Islam working together — is to observe and announce the appearance of the new moon that heralds the beginning of each *hegira* month.

A project that does have the go ahead is the construction of a facility for analysing, treating and printing the information that the KACST Directorate of Space Sciences can already receive from the Landsat and Spot satellites. The directorate is also to establish a laser ranging station.

Another source of some scientific funds in the country is the King Faisal Foundation which, for example, is establishing a cancer research centre. Only about a quarter of its SR240 million total disbursements to the end of 1985 were spent on scientific and medical research, however, and much of that outside Saudi Arabia. The rest has been spent mainly on the construction and administration of schools, colleges and mosques as well as on educational scholarships.

Set up by King Faisal's eight sons after his death in 1975, the foundation's funds are derived from investment income. An initial investment of £180 million derived from the King's estate in 1981 has grown to £220 million but the aim is to double it. Attempts to persuade individuals to contribute substantially have not been very successful, admits Prince Khaled Al Faisal, one of the late King's sons and director general of the foundation. Another activity of the foundation is the annual award of prizes in science and medicine (*Nature* 326, 118; 1987).

Saudi science is only at its beginning. Its aims are largely applied and its progress is still very dependent on collaborations with the non-Islamic world. But as national strength builds up, and as long as oil prices are at least reasonable, the enterprise seems set to grow rapidly. When it will mature is anybody's guess.

Peter Newmark

Knowledge-based prediction of protein structures and the design of novel molecules

T. L. Blundell, B. L. Sibanda, M. J. E. Sternberg & J. M. Thornton

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

Prediction of the tertiary structures of proteins may be carried out using a knowledge-based approach. This depends on identification of analogies in secondary structures, motifs, domains or ligand interactions between a protein to be modelled and those of known three-dimensional structures. Such techniques are of value in prediction of receptor structures to aid the design of drugs, herbicides or pesticides, antigens in vaccine design, and novel molecules in protein engineering.

TECHNOLOGICAL developments of industrial, clinical and agricultural importance may be achieved in the coming years by imitation of the interactions between macromolecules and ligands that occur naturally in the living cell. For example, the design of drugs, herbicides and pesticides may be improved from knowledge of the interaction of a molecule with an isolated receptor, enzyme or nucleic acid. The specificity, stability or activity of engineered hormones of clinical importance or enzymes of value to the chemical industry may be improved from knowledge of proteins in general¹. New peptide and protein vaccines may also be designed from information on antibody-antigen binding². In the longer term new molecular electronic devices may be constructed using novel molecules—perhaps proteins—that aggregate in a predetermined way (as they do in living cells) and provide a template for molecules that conduct electrons; these will be the new biological microchips³.

All these processes require information on the shape, charge, chemical function and dynamical flexibility of at least two interacting molecules. In most cases, at least one of the molecules will be a protein—enzyme, receptor, immunoglobulin, redox protein or polypeptide hormone. Much depends on knowledge of detailed three-dimensional structures defined by X-ray analysis, although medium resolution structures of small proteins in non-crystalline environments will increasingly come from two-dimensional NMR techniques⁴ in the future. However, our knowledge of three-dimensional structures of proteins defined by these methods has increased only slowly, in spite of the greater number of laboratories involved in this work.

In contrast, the advent of recombinant DNA techniques has led to an explosion of information concerning the sequences of receptors and enzymes important for drug, herbicide and pesticide design. At the same time site-specific and deletion mutagenesis offer new possibilities of engineering new proteins with point mutations, with hybrid domains and even with new functions⁵. Very few proteins that have been sequenced and can now be expressed and manipulated in the laboratory have known three-dimensional structures, and so rational approaches to design are difficult. However, computer technology has also developed at a rapid rate and offers some compensatory possibilities. These include the use of database technology to store, retrieve and compare the known sequences; computer graphics to display models and manipulate known three-dimensional structures⁶; and computer simulation to calculate low-energy conformers, normal modes and molecular dynamics⁷. Together these new technologies offer the chance of exploiting our experimental knowledge of the three-dimensional structures of proteins in a rational approach to the modelling of protein interactions and the design of novel molecules.

In this review we emphasize our own view that modelling of proteins and their interactions is most usefully carried out using a knowledge-based approach. This depends on the identification of analogies between a protein that is to be modelled and other proteins of known three-dimensional structure at all levels in the hierarchy of protein organization: secondary structure, motifs, domains and quaternary or ligand interactions.

Sequence alignment

The first step in any modelling is to convert the DNA sequence available into the primary structure of the gene product and search for sequence homologies with known sequences of other proteins. Information on DNA sequences is available in databanks such as those of the European Molecular Biology Laboratory (EMBL), Heidelberg, of the National Biomedical Research Foundation at Maryland, United States, or GenBank (Bolt, Beranach and Newman Inc.) (4,000,000 bases and 5,000 entries). Protein sequences are collected together in the Protein Information Resource (PIR) databank at National Biomedical Research Foundation, Maryland, or the NEWAT databank in the USA. The PIR databank now contains over 3,000 protein sequence entries, many of which have been derived by translation of DNA sequences (for a review of sequence databases, see ref. 8). The National Biomedical Research Foundation provides some software to search their databanks, to compare user-specific segments with segments of the same length in the database, to align two sequences, to detect similarities and to score and display the degree of similarity.

The sequence alignment algorithms of Needleman and Wunsch⁹, Sellers¹⁰ and Waterman *et al.*¹¹ use dynamic programming methods to carry out a global comparison of two sequences. Their success depends very much on the degree of similarity and may give variable results depending on the gap-penalty parameters chosen, even for closely related sequences^{12,13}. For sequences that are similar (>25%) such automatic procedures will identify the homology above the background of randomized sequences. As an approximate guide, if the alignment score is more than six standard deviations above that for random alignment, then most residues in secondary structures will be correctly aligned¹³. Because insertions and deletions often occur at the loop regions of proteins between secondary structures, improvements can be made by introducing penalties for insertions/deletions in α -helices or β -strands^{13,14}.

For very distantly related sequences, homology may be restricted to a few key residues or sequence segments whose separation along the chain may vary considerably between proteins. There have been several developments of algorithms to identify such local homologies^{15–18}.

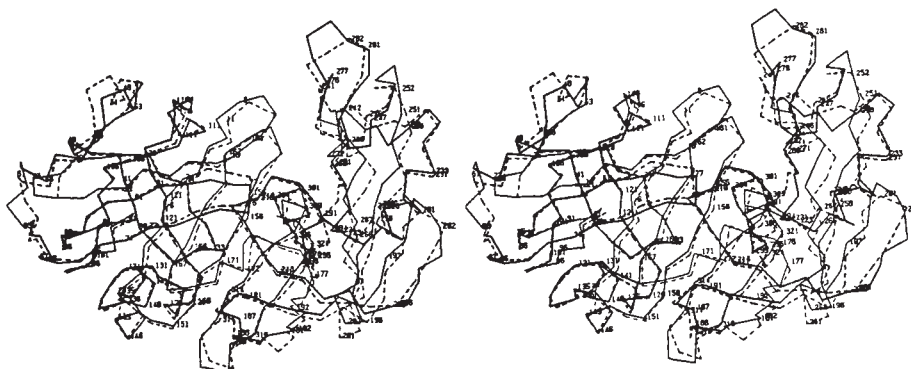


Fig. 1 A stereo view comparing the three-dimensional structures of two aspartic proteinases, endothiapepsin and penicillopepsin, indicates that the central cores, and active-site-cleft residues are closely similar. Diversity is achieved mainly in the loop regions that occur at the periphery of the bilobal enzymes.

The tertiary structures of homologous proteins

Comparison of the tertiary structures of homologous proteins shows that three-dimensional structures are conserved in evolution more than protein primary structures and considerably more than DNA sequences (ref. 19 for review). In recently diverged molecules the elements of secondary structure— α -helices and β -strands—are arranged in closely comparable three-dimensional topologies. Amino-acid replacements occur most often in surface positions so that the main chain conformations are little affected²⁰. More radical insertions and deletions tend to occur in surface loops between the secondary structure units although insertions are allowed in some β -strands to give rise to a β -bulge²¹. In some families of proteins such as the insulin family this divergence has occurred with complete conservation of the hydrophobic core²².

However, in most protein families divergence is accompanied by changes in the hydrophobic core. This has been examined in an elegant series of papers by Chothia and Lesk^{23–25} who have shown that although the volume of the core remains approximately constant in many families, amino-acid replacements of hydrophobic core residues are usually accommodated not by complementary amino-acid substitutions but rather by small shifts of secondary structural elements. They have shown that in the α -helical globins large relative shifts of certain helices are observed, but in the β -sheet family of azurins/plastocyanins relative rotations of the β -strands have occurred. To a certain extent this appears to be family dependent and is less evident not only in the immunoglobulins where an intersheet disulphide bridge pins them together but also in the core of other β -sheet proteins such as aspartic proteinases (Fig. 1).

What is the extent of such differences in the hydrophobic cores? To define this quantitatively, the topological equivalence of residues in homologous proteins is identified by a least-squares procedure that fits the proteins in three dimensions and calculates the root-mean-square (r.m.s.) differences of equivalent C_α atoms. A problem arises in the definition of which residues constitute the core. Chothia and Lesk²⁶ first superposed the major elements of secondary structure and then extended them to include additional residues at each end so long as the deviations did not exceed 3 Å. Using this definition the 'common core' included over 90% of the amino acids for proteins of >50% identity. At 50% identity the r.m.s. deviation averaged ~1 Å but there was a correlation between r.m.s. deviation and percentage residue identity of the complete protein.

This definition of the common core may not always be most appropriate. If the core is defined by those residues whose side chains are inaccessible to solvent, a similar correlation is found²⁷ although the number of residues and the average r.m.s. deviations are much smaller (Fig. 2).

Some of the differences between homologous structures arise from crystallographic errors and from conformational differences in the different crystalline environments. This can be estimated from comparison of proteins of identical sequence whose tertiary structures have been defined independently. The

hydrophobic cores may differ by only 0.2 Å in well defined, high resolution (<1.8 Å) structures but for all residues the differences are likely to be 0.3–0.6 Å. Most significantly the differences can be 1 Å for structures at intermediate (2.5–3.0 Å) resolutions.

Modelling by homology

Given homologous proteins with three-dimensional structures defined by high resolution X-ray analysis, how should we proceed to construct a model? Historically, this has been achieved in several simple stages starting with alignment of the sequences, followed by creation of insertions, deletions and replacements in the known three-dimensional structure of the homologous protein. This was first carried out in 1969 by Browne and co-workers²⁸ in the construction of a model of α -lactalbumin based on the three-dimensional structure of lysozyme which had been defined by X-ray analysis. Although it was achieved using physical models, subsequent attempts exploited computer graphics as the technology developed, especially the program FRODO²⁹. The initial models have been refined using energy minimization techniques to give final structures without steric clashes³⁰. Molecules modelled in this way have included

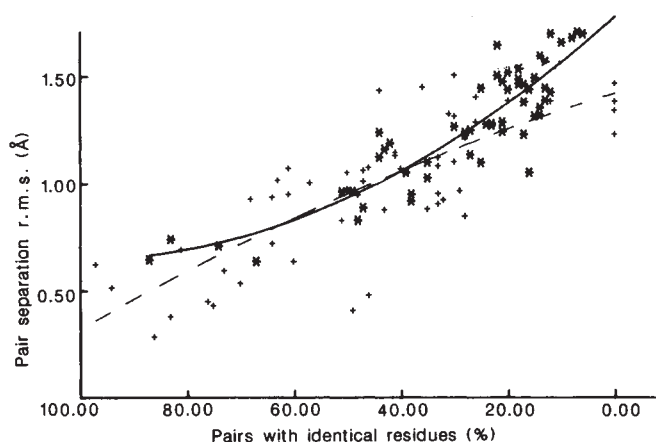


Fig. 2 A graph in which stars (*) indicate the root-mean-square (r.m.s.) separation of topologically equivalent residue (C_α) positions in pairs of proteins plotted as a function of the degree of identity of equivalent residues that are <3 Å in separation when the two proteins are arranged as a best fit. The best fitting curve is given as a continuous line. Crosses (+) show residues whose side chains are <7% accessible to the solvent and the best-fitting curve is given by a dashed line. The percentage identities refer only to the residues compared. For any particular protein, the percentage homology is increased but the r.m.s. separation and the number of residues are decreased when the solvent-inaccessible residues are compared. However, apart from proteins of low homology, the relationship between the r.m.s. separation and sequence homology is similar in each analysis (T. Hubbard, personal communication).

relaxins^{31,32}, insulin-like growth factors³³, serine proteinases²⁰, HLA-DR antigens³⁴, angiogenin³⁰, aspartic proteinases such as renin³⁵⁻³⁸ and immunoglobulins^{39,40} (ref. 41 for review).

Modelling using multiple structures

With over 300 structures in the Brookhaven databank, of which only ~100 are non-homologous, there may well be more than one structure available to serve as a basis for modelling. One approach to this problem⁴² is to model the structure on the basis of each available known tertiary structure and to test the resulting models for packing of side chains and their solvent accessibilities. Thus, calmodulin was constructed on the basis of intestinal calcium-binding protein and carp parvalbumin⁴². Although each sequence had approximately the same sequence homology with calmodulin, certain features of the model based on intestinal calcium-binding protein led to this being the preferred model.

An alternative knowledge-based approach is to use simultaneously but selectively the information from all the known tertiary structures of the homologous family. For example there are five high-resolution mammalian serine proteinase structures defined by X-ray analyses. The first step is to align the sequences and tertiary structures of the homologous proteins taking care that all structures are evenly weighted; it is not satisfactory to fit by least squares criteria all the tertiary structures to one of the family. One approach is to construct a 'framework' that contains a virtual atom at points in three-dimensional space of the topologically equivalent residues common to the family⁴³. The second step is to align the new sequence with those of the family so that topologically equivalent residues of the 'framework' are equivalent in the sequence alignment. This generally involves some realignment from comparisons of the sequence alone; it may be achieved by hand or by using the template sequence routines of Taylor¹⁸. Fragments of each of the homologous proteins which are closest in sequence are selected; generally it is preferable to select fragments which join one secondary structure to another and if possible link one element of the framework to the other. For the modelling of the serine protease domain of tissue plasminogen activator on the basis of the then-available trypsin, elastase, chymotrypsin and kallikrein structures, seventeen fragments were used (Fig. 3). A similar procedure has been developed independently by Levitt and co-workers³⁹.

Insertions and deletions in loop regions

Because the majority of significant differences in protein structure occur in loop regions, these are the most difficult to construct. The choice of a conformation found in a loop of equivalent length in a homologous protein is often a good guide, even if the sequence differs²⁰, although this will occasionally be misleading^{44,45}. Nevertheless, the problem remains how to identify loop sequences of the same length where conformations differ. If there is no sequence of equivalent length in any homologous protein, conformations from other proteins in similar supersecondary structures can be explored¹.

The database for such prediction of conformation in loops is now being constructed. Sibanda and Thornton⁴⁶ and Milner-White and Poet⁴⁹ have detailed the conformations of β -hairpin structures (loops between two adjacent antiparallel β -strands) and similar analyses are proceeding for loops between two α -helices and between α -helices and β -strands⁴⁷. Detailed analysis of β -hairpin loops shows that in short loops certain structures with characteristic sequences recur many times in unrelated proteins (ref. 46; Fig. 4 and Table 1). In general there are few longer loops; the exceptions often include prolines and the sequence Pro-Gly is common.

Although there is little sequence homology between loops, certain glycines are often conserved and there is a preference at certain positions for residues that can help form β -strands or α -helices. For modelling we can then adopt a systematic



Fig. 3 A model of tissue plasminogen activator (*f*) constructed on the basis of four other mammalian serine proteases of defined three-dimensional structure. Seventeen fragments are selected on the basis of sequence homology. The first in the sequence is shown in *a* and *b*, *c* and *d* show the addition of subsequent fragments as dotted lines: *e* and *f* show the last two fragments in the sequence. Four further fragments are added from other proteins (D. Athwal, and T. Harris personal communication; T.L.B., unpublished data).

approach that involves first identifying the loop length, secondly selecting a conformer on the basis of sequence and thirdly testing the conformer by least-squares fitting to the β -strands of the model framework. Similar techniques are applicable in principle for modelling all types of loops.

The power of this technique is shown in models of mouse and human renins^{35,36} and chymosin⁵⁰ that were modelled on the basis of endothiapepsin. For example, sequences of equivalent loops in endothiapepsin and chymosin are given as:

Residue number	197	198	199	200	201	202	203	204
Endothiapepsin	Y	A	V	<u>G</u>	<u>S</u>	G	T	F
Chymosin	V	T	I	<u>S</u>	<u>G</u>	V	V	V

where the turns are underlined. In endothiapepsin this is a two residue type II' turn with glycine at the first position. The sequence of chymosin with a glycine at the second position indicates that this is most probably a type I' turn. In a second example a deletion of one residue is observed. Thus:

Residue number	238	238A	239	240	241	242	243	244	245	246	247
Endothiapepsin	A	K	S	S	<u>S</u>	<u>S</u>	<u>V</u>	<u>G</u>	G	Y	V
Chymosin	A	T	Q	<u>N</u>	<u>Q</u>	<u>Y</u>	—	<u>G</u>	<u>E</u>	F	D

In endothiapepsin this is a standard four-residue loop with a glycine at position 4. In chymosin there is a deletion of one residue leading to a three-residue loop, but the presence of glycine implies that this may lose one hydrogen bond to give a standard five-residue loop with a glycine at position 4 (Fig. 5).

A similar approach has been adopted by Jones and Sirup⁵¹ in their extension of FRODO. They search a structural database to find all possible loops of the correct length using an elegant method based on the distance between residues displayed as a matrix. A computer search for suitable motifs involves pattern matching of distance matrices. Alternatively the appropriate sequences and conformations can be identified in a relational database¹.

Where no structure appropriate for the modelling exists in conformations of proteins defined by high resolution X-ray analysis, we must adopt an alternative *ab initio* approach such

Table 1 Systematic modelling of β -hairpins

SYSTEMATIC MODELLING OF β -HAIRPINS

← REPLACEMENT →

SET	DOUBLE H-BOND			SINGLE H-BOND		ALTERNATIVES
Deletion ↑ ↓ Insertion A + 1 B + 1 C + 2 D + 3	2:2 	• Type I' Gly - Asn - Gly - Asp α_L G	• TYPE II' - Gly - Ser - Thr - G α_R	TYPE I - X - X - α_R α_R	2:4 	unusual - various 6:6 6:8 10:10 10:12
	3:3 	Rare - various			3:5 	• TYPE I [1-4] + Gl β -bulge - X - X - X - Gly - X β α_R α_R G β 7:7 7:9 11:11 11:13
	4:4 	• Type I [1-4] α_R α_R α_R α_L - X - X - X - Gly	Various		4:6 	Many different conformations. 8:8 8:10 12:12 12:14
	5:5 	Many different conformations.			5:7 	Many different conformations. 9:9 9:11 13:13 13:15

A schematic representation of possible β -hairpin structures found in proteins defined by X-ray analysis. A similar analysis has been carried out by Milner-White and Poet⁴⁹. Structures in horizontal lines (such as 2:2 and 2:4) have identical numbers of residues and represent differing conformations that may be preferred by particular amino acid sequences; these are useful for modelling replacements. Structures in columns represent hairpin conformations with differing numbers of residues and these are used in modelling insertions or deletions. Loss of the top hydrogen bond implies a horizontal movement to the right (3:3 $\rightarrow$ 3:5) whereas loss or gain of two hydrogen bonds implies a change such as 3:5 $\rightarrow$ 7:9 which belongs to the same class. Preferred sequence patterns are indicated (from ref. 46 and B.L.S., J.M. and T.L.B. in preparation). Loop lengths are defined by two numbers X; Y where X is the number of residues not involved in β -strands and Y is the number of residues that do not have both amide and carboxyl groups involved in H-bonds characteristic of β -strands.

as use of a distance geometry algorithm⁵² or a systematic search procedure of low energy conformers using molecular dynamics⁴⁸.

Having created the appropriate backbone, the amino acids that differ between known structures and the modelled structure must be replaced. For this we assume that the side-chain torsion angles for equivalent bonds are the same and that the side chain occupies a similar region in three-dimensional space in the family of structures. A set of rules for replacement of residues based on conformations of side chains at topologically equivalent positions in homologous proteins (M. Sutcliffe, personal communication) has been derived.

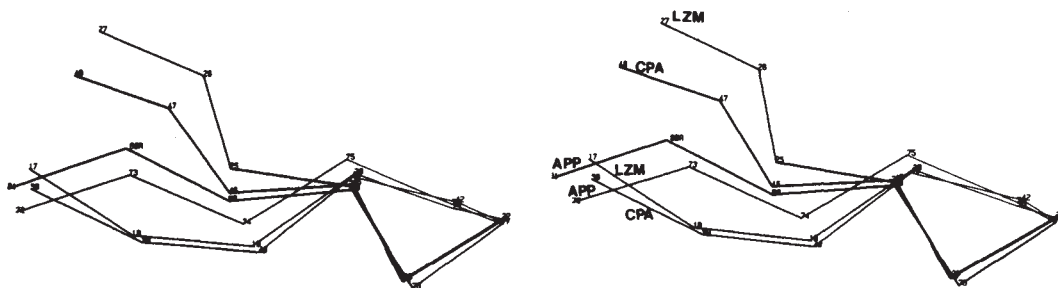
Energy minimization and molecular dynamics

Where the proteins have sequence homology of 50% or more, the models predicted by the methods described here will be probably correct to better than 1 Å although individual side chains may be more in error. Some but not all the errors can be removed by minimization of the potential energy using standard programs such as AMBER⁵³ or CHARMM^{54,55}. Whereas X-ray analysis gives a time- and space-averaged structure in a

crystalline lattice with perhaps 40% solvent, energy minimization gives the structure at a single minimum usually *in vacuo* as solvent is not often simulated. Efforts to simulate the aqueous environment using Monte Carlo methods and perhaps to model specifically tightly bound waters⁵⁶ are particularly helpful for surface side chains although some contraction of the protein volumes tends to occur⁵⁷ unless the molecule and solvent are included in a repeating lattice. Use of energy minimization in this way finds only a local minimum and may be expected to be useful if the errors are relatively small (usually <1 Å).

To explore local changes requiring larger shifts, methods of constraining the major part of the structure whilst allowing freedom in the region of interest have been explored⁵⁵. With the extent of the calculation reduced, side-chain torsion angles can be varied to locate conformations for subsequent energy minimization. Alternatively, energy calculations or molecular dynamics may often be useful to explore a range of local conformations, as developed by Robson and co-workers⁵⁸. Energy minimization used in the conventional way will not be useful as a local energy minimum can always be found. No one has modelled successfully changes such as those identified by Lesk

Fig. 4 A stereo view of three similar β -hairpin structures (of type 3:5, see Table 1) superposed with r.m.s. error (calculated for the five residues in the loop) of 0.56 Å. The examples given are from penicillopepsin (APP: residues 72–82), carboxypeptidase A (CPA: residues 38–48) and lysozyme T_4 (LZM: residues 17–27).



and Chothia²³ in the globins where an α helix can differ in position by 7 Å and rotational angle by 30°. The best chance appears to be in a simplified molecular dynamics procedure with rigid secondary structural units, or an interactive docking procedure such as that of Wodak *et al.*⁵⁷.

How correct are the models?

Several modelled proteins have been defined subsequently by X-ray analysis. The first of these models, that of α -lytic proteinase, was flawed by an incorrect sequence alignment. Others have been more successful. For example, the three-dimensional structures of rat γ 3-1 crystallin and the orthologous protein γ IV from calf were modelled on the basis of γ II-crystallin⁵⁹. The three-dimensional structure of the calf γ IV defined by X-ray analysis at 2.3 Å resolution has been shown to have a r.m.s. difference of ~0.7 Å with the model overall and 0.5 Å in each domain⁶⁰. Chothia *et al.*³⁹ have modelled the variable domain of immunoglobulin and have correlated their prediction with the recently defined X-ray analysis of Poljak and co-workers⁶¹. They have not only predicted the loop conformations correctly in 4 of the 6 hypervariable regions, but also attained a r.m.s. difference of 1 Å in the framework region for the main chain.

When a structure is modelled on the basis of one homologous protein, the relationship of the percentage sequence identity to the r.m.s. difference given by Chothia and Lesk is a reasonable guide to the precision. A model of sperm whale myoglobin based on either human or equine β -chains of haemoglobin has an r.m.s. difference of 1.40 Å from the structure defined by X-ray analysis. The precision can be improved by fitting the modelled

structure to the framework derived from all the homologous structures. This leads to a smaller r.m.s. difference of 1.25 Å between the model and the X-ray structure⁴³ when the α - and β -chains of equine and human haemoglobins are used.

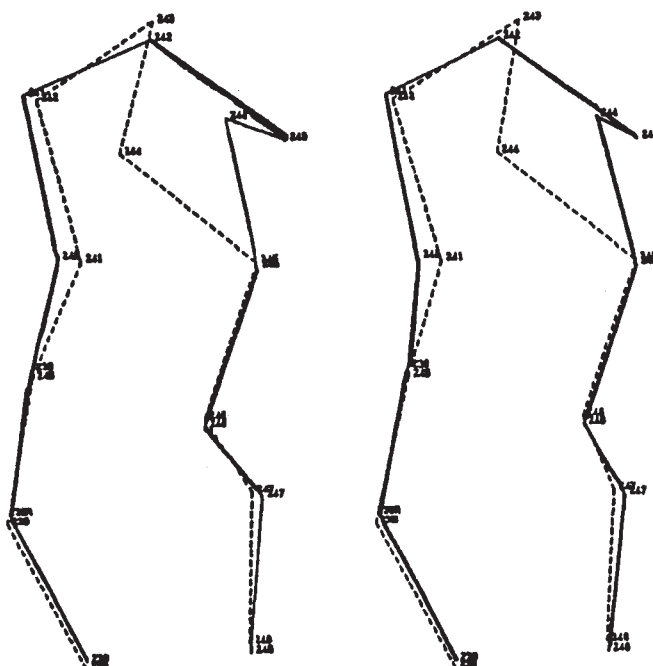
Although the absolute value of potential energy of the model as calculated by programs such as CHARMM is not a good guide to its correctness⁶², the distribution of the hydrophobic side chains and the nature of the solvent-accessible surface are more sensitive indicators; these are features that can be quantified but are also easily assessed visually. For example, a buried charged side chain that is not hydrogen-bonded is almost certainly an indication that the model is incorrect.

Perhaps the most useful indication of the correctness of a model is its ability to predict a molecular interaction that can be experimentally tested. Models of insulin-like growth factors³³ demonstrated that these molecules might bind insulin receptors. Electrostatic potential surfaces generated for the model of calmodulin have suggested a binding site consistent with experimental data for basic amphiphilic α -helical peptides⁴². The size and chemical nature of the predicted combining sites of anti-lysozyme monoclonal antibodies have been tested and are consistent with epitope boundaries⁴⁰. The size, shape and position of the specificity pockets of human renin^{36–38} have been assessed as useful predictors for the design of inhibitors.

Future developments

There are two challenges for the future. First, to extend the method to cases where no obvious sequence homology may be found, but where the unknown structure is probably constructed

Fig. 5 A model of chymosin residues 238–248 which form a β -hairpin (broken line). The model is based on the high-resolution refined structure of endothiapepsin (continuous line). A single residue deletion in the chymosin structure results in a remodelling of the family 4:4 β -hairpin of endothiapepsin as a family 3:5 hairpin in chymosin.



from the structural motifs, $\beta\alpha\beta$ units, $\beta\beta$ hairpins, an 8-fold $\alpha\beta$ barrel such as found in triose phosphate isomerase⁶³ or a jelly-roll structure, such as is found in the viruses⁶⁴. The problem is to identify these motifs from the sequence through detailed study of the critical sequence patterns for each motif and to establish the sequence fingerprint against which a sequence of unknown structure can be compared.

For example, a particular pattern of a glycine and a serine along with a few conservatively varied residues and hydrophobic regions in a sequence of ~40 residues has been used to identify a Greek key structure in the surface protein S of the sporulating bacterium *Myxococcus xanthus* even though there is no significant overall homology⁶⁵. Another pattern Gly-X-Gly-X-X-Gly, where X is any amino acid, has been useful for identification of an α/β structure in the p21 protein⁶⁶ and in tyrosyl kinases⁶⁷.

Second, we must use the methods to design novel molecules. For protein engineering replacements, small insertions and deletions, the approach is a natural extension of the modelling procedures. Sequences in similar structural motifs can be identified from known protein structures and grafted onto the known structure. Local conformations can be explored by global searches, restrained minimization approaches or molecular dynamics followed by energy minimization. In the design of hybrid molecules for example linking monoclonal antibody Fab fragments to enzymes such as proteases for dissolving blood-clots, or to toxins for killing tumour cells, the identification of stable domains and the design of linker regions will be of central importance⁶⁸. In the design of novel tertiary structures which might be endowed with special binding or catalytic functions, the complete technology evolved for modelling will be essential.

In drug design the modelling of the receptor must be comple-

mented with the definition of the ligand-binding interaction. Much is to be learnt from interactions in homologous systems such as other members of the enzyme family. Thus, a starting point for modelling human angiotensinogen interactions with renin is usefully obtained from high resolution X-ray studies of inhibitors with other aspartic proteinases such as endothiapepsin^{1,69}. Also, detailed analyses of side chain-side chain and side chain-ligand interactions can define preferred orientations, for example of phenylalanine rings with oxygens⁷⁰, sulphur⁷¹ and other aromatic groups⁷².

For vaccine design, the synthetic molecule must mimic the antigen to stimulate the appropriate immune response^{73,74}. Optimal activity will be achieved by designing the synthetic molecule to mimic the critical features of the native antigen as closely as possible. As in drug design it may be advantageous to restrict flexibility, for example by ring closure. Alternatively, an antigenic loop peptide may be grafted from one protein into another using the technique of protein engineering. Indeed it may be possible to produce a protein vaccine that incorporates antigenic peptides from many proteins and confers multi-valent immunity. For rational design of such complex vaccines and even simple epitopes, model building using powerful computer graphics including superposing, annealing and energy minimizing structures, will provide starting points to guide the experimenter.

We thank Willie Taylor, Ian Tickle, Suhail Islam, Michael Sutcliffe, Tim Hubbard, David Barlow, Dee Atwal, Ilyas Haneef, Andrew Hemmings, James Milner-White, Geoff Barton, Mark Edwards, Marketa Zvelebil, Jus Singh and Stephen Bryant for discussion of the ideas discussed in this review. We thank the SERC, ICI, CellTech, Glaxo and Sturge for financial support.

- Blundell, T. L. *et al. Phil. Trans. R. Soc. A* **317**, 333-344 (1986).
- Porter, R. & Whelan, J. (eds) *Synthetic Peptides as Antigens, CIBA Foundation Symposium* 119 (Wiley, Avon, 1986).
- Haddon, R. C. & Lamola, A. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1874-1878 (1985).
- Wagner, G. & Wuthrich, K. *J. molec. Biol.* **155**, 347-366 (1982).
- Winter, G. & Fersht, A. R. *Trends Biotechnol.* **2**, 115-119 (1984).
- Richards, W. G. (ed.) *J. molec. Graphics* **4**, 1-73 (1986).
- Brooks, B. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6571-6575 (1983).
- Kneale, G. G. & Bishop, M. *J. Cabios Rev.* **1**, 11-17 (1985).
- Needleman, S. B. & Wunsch, C. D. *J. molec. Biol.* **48**, 443-453 (1970).
- Sellers, P. *J. appl. Math.* **26**, 787-793 (1974).
- Waterman, M. S., Smith, T. F. & Beyer, W. A. *Adv. Math.* **20**, 367-387 (1976).
- Fitch, W. M. & Smith, T. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1382-2386 (1983).
- Barton, G. & Sternberg, M. J. E. *Protein Engng* (in the press).
- Lesk, A., Levitt, M. & Chothia, C. *Protein Engng* **1**, 77-78 (1987).
- Goood, W. B. & Kanehisa, M. I. *Nucleic Acids Res.* **10**, 247-263 (1982).
- Sellers, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3041-3044 (1979).
- Boswell, D. R. & McLachlan, A. D. *Nucleic Acids Res.* **12**, 457-464 (1984).
- Taylor, W. R. *J. molec. Biol.* **188**, 233-258 (1986).
- Bajaj, M. & Blundell, T. L. *Rev. Biophys. Bioengng* **13**, 453-492 (1984).
- Greer, J. *J. molec. Biol.* **153**, 1027-1042 (1981).
- Richardson, J. S., Getzoff, E. D. & Richardson, D. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2574-2578 (1978).
- Cutfield, J. F. *et al. in Structural Studies on Molecules of Biological Interest* (ed. Dodson, G. C.) 501-508 (Clarendon Press, Oxford, 1981).
- Lesk, A. M. & Chothia, C. *J. molec. Biol.* **136**, 225-270 (1980).
- Lesk, A. M. & Chothia, C. *J. molec. Biol.* **160**, 325-342 (1982).
- Chothia, C. & Lesk, A. M. *J. molec. Biol.* **160**, 309-323 (1982).
- Chothia, C. & Lesk, A. M. *EMBO J.* **5**, 823-826 (1986).
- Hubbard, T. & Blundell, T. L. *Prot. Engng* (submitted).
- Browne, W. J. *et al. J. molec. Biol.* **120**, 97-120 (1969).
- Jones, A. T. *J. appl. Cryst.* **11**, 268-272 (1978).
- Palmer, K. A., Scheraga, H. A., Riordan, J. F. & Vallee, R. Z. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1965-1969 (1986).
- Isaacs, N. *et al. Nature* **271**, 278-281 (1978).
- Bedarkar, S. *et al. Nature* **270**, 449-451 (1977).
- Blundell, T. L., Bedarkar, S. & Humbel, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 180-184 (1978).
- Travers, P. *et al. Nature* **310**, 235-278 (1984).
- Blundell, T. L., Sibanda, B. L. & Pearl, L. H. *Nature* **304**, 273-275 (1983).
- Sibanda, B. L. *et al. FEBS Lett.* **174**, 103-111 (1984).
- Carlson, W., Karplus, M. & Haber, E. *Hypertension* **7**, 13-26 (1985).
- Akahone, P. *et al. Hypertension* **7**, 3-12 (1985).
- Chothia, C. *et al. Science* **233**, 755-758 (1986).
- La Paz, P., Sutton, B. R., Darsley, M. J. & Rees, A. R. *EMBO J.* **5**, 415-425 (1986).
- Ripka, W. C. *Nature* **321**, 93-94 (1986).
- O'Neil, K. & De Grado, W. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4954-4958 (1985).
- Sutcliffe, M. & Haneef, I. *Inf. Q. Prot. Cryst.* **18**, 11-18 (1986).
- Read, R. J., Fujiwara, M., Sielecki, A. R. & James, M. N. G. *Biochemistry* **22**, 4420-4433 (1983).
- Read, R. J., Brayer, G. D., Jarasek, L. & James, M. N. G. *Biochemistry* **23**, 6570-6575 (1984).
- Sibanda, B. L. & Thornton, J. M. *Nature* **316**, 170-174 (1985).
- Edwards, M., Sternberg, M. J. E. & Thornton, J. M. *Prot. Engng* (in the press).
- Moult, J. & James, M. N. *Proteins* **2**, 146-163 (1986).
- Milner-White, J. & Poet, R. *Biochem. J.* **240**, 289-292 (1986).
- Sibanda, B. L. thesis, Univ. London (1986).
- Jones, T. A. & Sirup, T. *EMBO J.* **5**, 819-822 (1986).
- Crippen, G. M. in *Chemometric Research Studies Series Vol. 1* (ed. Bandan, D.) 1-58 (Wiley, New York, 1981).
- Weiner, S. J. *et al. J. Am. chem. Soc.* **106**, 765-784 (1984).
- Brooks, B. R. *J. Computation Chem.* **4**, 187-217 (1983).
- Shih, H. L., Brady, J. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1697-1700 (1985).
- Hagler, A. T. & Moult, J. *Nature* **272**, 222-227 (1979).
- Wodak, S. J., Alard, P., Delhouse, P. & Renneboog-Squibin, C., *J. molec. Biol.* **181**, 317-322 (1984).
- Robson, B. & Platt, E. *J. molec. Biol.* **188**, 259-281.
- Summers, L. D. *et al. Expl Eye Res.* **43**, 77-92 (1986).
- Driessen, H. P. C. & White, H. in *Molecular Replacement* (ed. Machin, P. A.) 27-32 (Daresbury Laboratory, Daresbury, 1985).
- Amit, A. G. *et al. Science* **233**, 747-753 (1986).
- Novotny, J., Bruccoleri, R. & Karplus, M. *J. molec. Biol.* **177**, 787-818 (1984).
- Banner, E. *et al. Nature* **255**, 609-614 (1975).
- Abad-Zapatero, C. *et al. Nature* **286**, 33-39 (1980).
- Wistow, G., Summers, L. J. & Blundell, T. L. *Nature* **315**, 771-773 (1985).
- Wieranga, R. K. & Hol, W. G. *Nature* **302**, 842-844 (1983).
- Sternberg, M. J. E. & Taylor, W. R. *FEBS Lett.* **175**, 387-392 (1984).
- Neuberger, M. S., Williams, G. T. & Fox, R. O. *Nature* **312**, 604-608 (1984).
- Hemmings, A., Foundling, S., Sibanda, B. L., Wood, S. P. & Blundell, T. L. *Trans. Biochem. Soc.* **13**, 1036-1041 (1985).
- Thomas, K. A., Smith, G. M., Thomas, T. J. & Feldmann, R. J., *Proc. natn. Acad. Sci.* **79**, 4843-4847 (1982).
- Reid, K. S. C., Lindley, P. F. & Thornton, J. M. *FEBS Lett.* **190**, 209-213 (1985).
- Singh, J. & Thornton, J. M. *FEBS Lett.* **191**, 1-6 (1985).
- Nussenzweig, V. & Nussenzweig, R. in *Ciba Foundation Symposium* 119 (eds Porter, R. & Whelan, J.) 150-163 (Wiley, Avon, 1986).
- Geysen, H. M., Rodda, S. J. & Mason, T. J. in *Ciba Foundation Symposium* 119 (eds Porter, R. & Whelan, J.) 130-149 (Wiley, Avon, 1986).

Timing and site of nerve growth factor synthesis in developing skin in relation to innervation and expression of the receptor

Alun M. Davies*, Christine Bandtlow†, Rolf Heumann†, Sigrun Korsching†, Hermann Rohrer† & Hans Thoenen†

* Department of Anatomy, St George's Hospital Medical School, Cranmer Terrace, Tooting, London SW17 0RE, UK

† Department of Neurochemistry, Max-Planck-Institute for Psychiatry, 8033 Planegg-Martinsried, Munchen, FRG

We show that nerve growth factor (NGF) synthesis in developing skin begins with sensory innervation and that sensory neurons do not express NGF receptors until their fibres reach their cutaneous targets. Both cutaneous epithelium and mesenchyme synthesize NGF, the concentration of messenger RNA for NGF being higher in the more densely innervated epithelium.

A GIVEN target-field of the developing vertebrate nervous system is able to support the survival of only a limited number of neurons, superfluous neurons being lost in a phase of cell death that begins shortly after the start of target-field innervation¹. The proposal that neuronal survival during this critical phase of development depends on a supply of a neurotrophic factor from the target-field² has received substantial support from work on the protein nerve growth factor (NGF). NGF supports the survival of sympathetic and certain neural-crest-derived sensory neurons in culture³⁻⁶ and prevents loss of these neurons *in vivo* if administered during the period of neural death^{4,7,8}. Conversely, the presence of anti-NGF antiserum during development greatly increases the loss of these neurons^{3,4,9}. The amounts of NGF and messenger RNA encoding it in a variety of organs and tissues are correlated with the density of sympathetic innervation^{10,11}, and NGF is conveyed from the periphery by fast axonal transport to the perikarya of sympathetic¹² and sensory neurons¹³. The amino-acid sequence of NGF was determined over a decade ago¹⁴ and this information has been used recently for cloning the NGF gene^{15,16}. This has revealed the organization of the NGF gene^{15,16} and that the NGF molecule is derived from a long precursor protein^{15,17} by proteolytic cleavage.

Although extensive information is available on the biochemistry, molecular biology and physiological effects of NGF, a number of important questions remain unanswered: these include the identity of the cells that synthesize NGF in the developing target-field, the temporal relationship of NGF synthesis to innervation and the stage at which innervating neurons express NGF receptors and become dependent on NGF for survival. These issues are important not only for understanding how NGF regulates target-field innervation, but also for clarifying the function of NGF in development, as the findings of a number of studies suggest that NGF also attracts nerve fibres from a distance to their target-fields by chemotaxis¹⁸⁻²¹. To investigate these issues we have studied the mouse trigeminal ganglion and part of its cutaneous target-field (the maxillary process and the whisker pad that develops from it) at closely staged intervals before and after innervation (Fig. 1). We chose this cutaneous target-field because it is well circumscribed throughout development and is the most densely innervated cutaneous target-field in the mouse. In the mature whisker pad the epidermis receives a rich innervation from sensory endings that penetrate the underlying basal lamina to terminate on Merkel cells, and the dermis contains a variety of encapsulated and non-encapsulated sensory endings²²⁻²⁴.

In normal development of the mouse^{25,26}, nerve fibres emerge from the trigeminal ganglion at 9.5 days of gestation (denoted

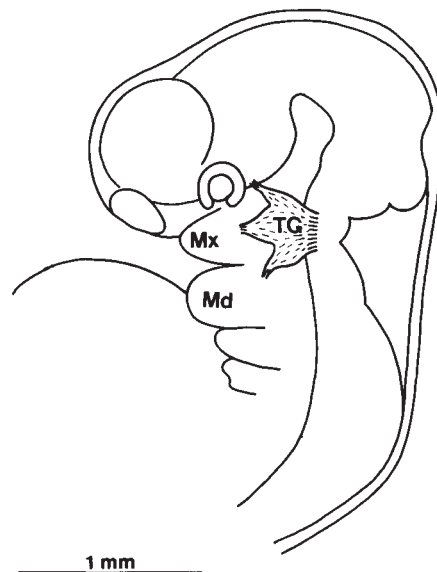


Fig. 1 Camera lucida drawing of the head of an E10 mouse embryo showing the location of the trigeminal ganglion (TG) and its peripheral target-field which largely consists of the maxillary process (Mx) and mandibular process (Md) of the first branchial arch.

E9.5, equivalent to Theiler stage 15) and increase in number to reach a maximum at E13. The earliest fibres reach the epithelium of the maxillary process by E11 and the latest reach their peripheral targets shortly after E15. Between E13 and birth (E19) over half the nerve fibres and neurons are lost as a result of cell death. In determining when and where NGF is synthesized and when its receptor is expressed, we have been able to clarify the role of NGF in development and show that it does not attract nerve fibres to their target-fields.

Changes in NGF level

Using a sensitive two-site immunoassay²⁷, NGF was not detectable in the maxillary process at either E10 or E10.5 (Fig. 2a). It was first detected at E11, the stage at which the earliest nerve fibres contact the target-field epithelium, and increased to reach a maximum level by E13, after which there was a fall. NGF was first detected in the trigeminal ganglion at E12, 24 hours after its appearance in the periphery, and increased in amount thereafter. This increase was due to axonal transport of NGF

from the periphery rather than local synthesis because NGF mRNA (below) was detected in the maxillary process but not in the ganglion. Thus, the fall in the amount of NGF in the target-field after E13 was due, at least in part, to its removal by axonal transport. It is of interest that the marked decrease in the concentration of NGF in the target-field that occurs after E13 (Fig. 2b) correlates in time with the onset of neuronal death in the trigeminal ganglion which also begins at E13.

Changes in the level of NGF mRNA

To determine the onset and time course of transcription of the NGF gene and the relationship of transcription to NGF synthesis, the level of NGF mRNA was measured at closely staged intervals throughout development by quantitative Northern blot analysis of total RNA using a ^{32}P -labelled, single-stranded complementary RNA probe¹⁰ (Fig. 3). The NGF message was not detectable in the maxillary process at E10. It was detectable in only very small amounts at E10.5, just 12 hours before NGF protein was first detected in the maxillary process. This close time sequence in the appearance of the NGF messenger and protein in the target-field, together with the parallel increase in the levels of these molecules during the early stages of innervation (Figs 2 and 3), indicate that the synthesis of NGF during development is governed by the level of its RNA.

Our finding that the start of target-field innervation and the onset of NGF mRNA transcription are closely related events in development raises the possibility that NGF synthesis is initiated by the arriving nerve fibres. Further experimental studies will be required to determine whether early nerve fibres induce NGF synthesis in virgin target tissue.

From its first appearance in the target-field, the total amount of NGF mRNA showed a continuous increase (Fig. 3b), suggesting that the overall production of NGF in the target-field increases throughout the phase of innervation. In contrast, the concentration of NGF mRNA increased to reach a peak at E12.5 after which there was a fall (Fig. 3c), suggesting that either individual cells produce less NGF, or NGF-producing cells are progressively outnumbered by non-producing cells as a result of target-field growth.

Site of NGF synthesis

To determine whether NGF synthesis occurs in either the epithelium (presumptive epidermis) or mesenchyme (presumptive dermis) of the target-field, these components were separated enzymatically and NGF mRNA was assayed in them (Fig. 4a). Both epithelium and mesenchyme contained NGF mRNA, and although the total amount in mesenchyme was somewhat greater than in epithelium throughout the period studied (E11 and E13), the concentration was 5–10-fold higher in epithelium owing to the small size of the latter (Fig. 4b). These findings suggest that NGF is synthesized in both the epithelium and mesenchyme of developing skin. The higher concentration of NGF mRNA in epithelial cells accords with the larger numbers of sensory endings that terminate in the epidermis compared with the dermis in mature skin²³.

To determine whether NGF synthesis is restricted to certain regions of the epithelium and mesenchyme, NGF mRNA was localized in tissue sections by *in situ* hybridization using an ^{35}S -labelled NGF cRNA probe²⁸ (Fig. 5). In agreement with our finding that the concentration of NGF messenger is higher in epithelium than mesenchyme, the epithelium was more heavily labelled than the mesenchyme in sections of the E13 whisker pad. The resolution of *in situ* hybridization is insufficient to identify unequivocally individual cells expressing a very rare messenger, so it has not been possible to ascertain whether all or only some epithelial cells synthesize NGF. Nonetheless, there were no obvious regional differences in the distribution of NGF mRNA in the epithelium; both the surface and follicular epithelium were heavily labelled. Although the mesenchyme of the whisker pad was labelled throughout, the region lying just

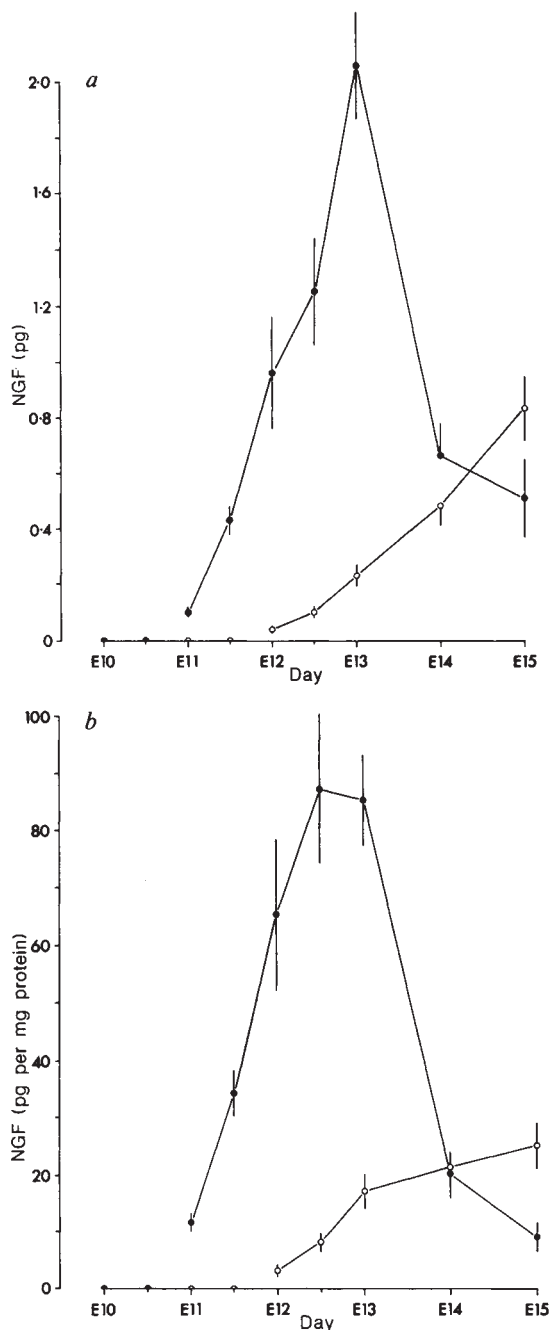


Fig. 2 Results of the two-site enzyme immunoassay for NGF²⁷ in tissues of E10 to E15 mouse embryos. **a**, Amount of NGF in pg per maxillary process/whisker pad (●) and trigeminal ganglion (○). **b**, Concentration of NGF in these tissues expressed in pg NGF per mg total protein (protein concentration was measured by the method of Lowry³³). The mean and s.e.m. of the three or more separate assays carried out at each age are plotted. To avoid errors in the detection of NGF in the tissues of younger embryos, large numbers of maxillary processes and ganglia (50–100) were used in each assay.

beneath the surface epithelium was more heavily labelled. Because this region gives rise to the dermis of the skin, the distribution of NGF mRNA in the mesenchyme accords with the termination of large numbers of nerve fibres in the dermis compared to subcutaneous tissues.

The widely accepted view that NGF is a target-cell-derived neurotrophic factor has recently been challenged by Rush and colleagues^{29,30} who, in a recent immunohistochemical study of the denervated iris, detected NGF in Schwann cells but not in

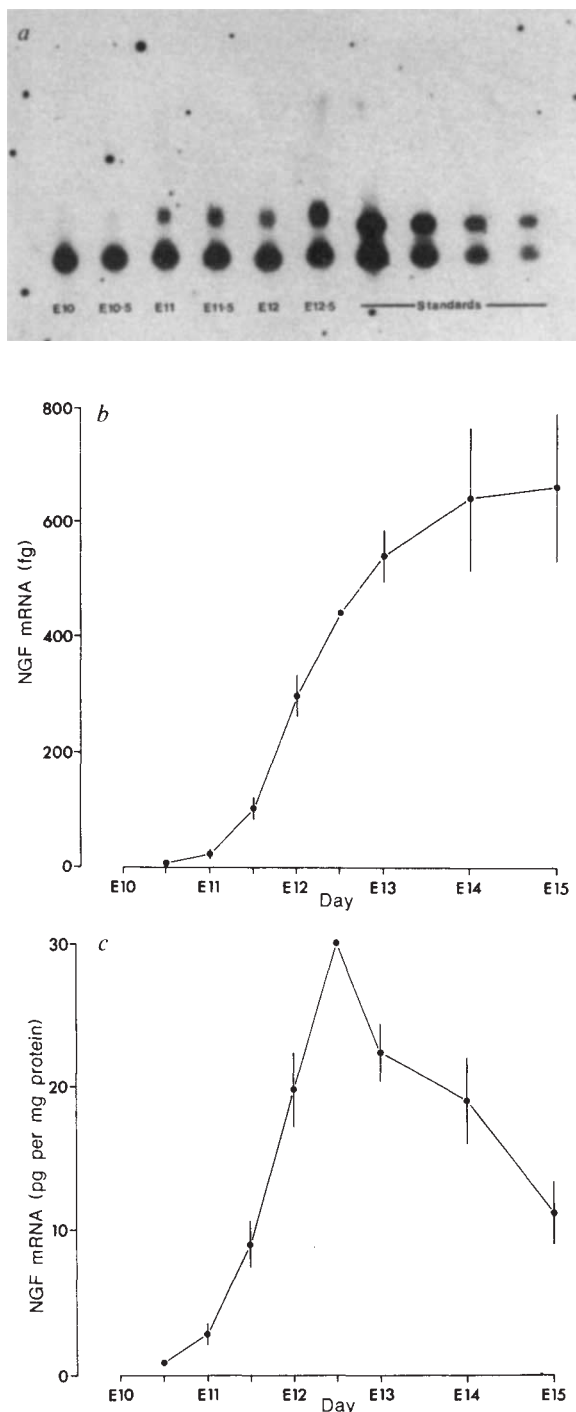


Fig. 3 Measurement of NGF mRNA in the maxillary process/whisker pad by quantitative Northern blot hybridization¹⁰. *a*, Autoradiograph of a Northern blot of tissue total RNA hybridized with a ³²P-labelled cRNA probe derived from the NGF cDNA clone by run-off transcription in an SP-64 vector. Total mRNA from E10 to E12.5 maxillary processes was run in the first six lanes. The upper band in these sample lanes, corresponding to 1,300 base pairs (bp), represents NGF mRNA and the lower band, corresponding to 510 bp, represents a short NGF RNA transcript which was added to tissue samples before extraction of RNA to monitor any losses of RNA during this step. The last four lanes represent known amounts of this short transcript and an NGF RNA transcript of similar length to NGF mRNA to calibrate the densitometric quantification of sample lane bands. *b*, Amount of NGF mRNA in fg per maxillary process/whisker pad determined by densitometry. *c*, Concentration of NGF mRNA expressed in fg per mg total protein. The mean and s.e.m. of the three or more separate assays carried out at each age are plotted.

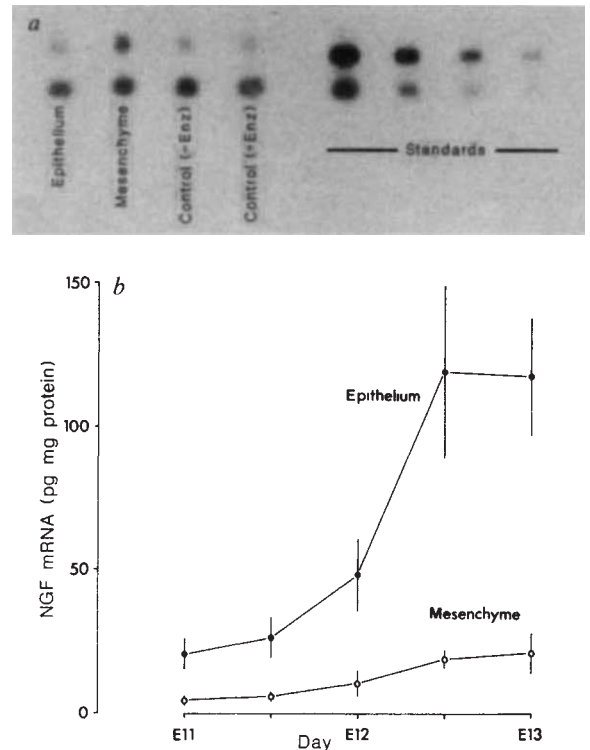


Fig. 4 Measurement of NGF mRNA in the isolated epithelium and mesenchyme of the maxillary process by quantitative Northern blot hybridization. *a*, Autoradiograph of a Northern blot of total RNA from E12 maxillary process tissue hybridized with the ³²P-labelled cRNA probe. Lanes 1 and 2; RNA extracted from the epithelium and mesenchyme of the same number of maxillary processes; lanes 3 and 4, controls to ascertain whether collagenase treatment interferes with detection of NGF mRNA; lane 3, RNA from a number of whole maxillary processes; lane 4, RNA from the same number of maxillary processes exposed to enzyme but not separated into their component tissues. A similar amount of NGF mRNA was detected in these lanes indicating that enzyme treatment has no effect on the detection of messenger. *b*, Concentration of NGF mRNA expressed in fg per mg protein in the epithelium and mesenchyme of the E11 to E13 maxillary process. The mean and s.e.m. of the three or more separate assays carried out at each age are plotted.

Methods. Maxillary processes were separated into component epithelium and mesenchyme by treatment with 0.5 mg ml⁻¹ collagenase (Worthington) in Hanks balanced salt solution for 30 min at 4°C. This treatment preferentially digests the basal lamina, allowing the sheet of epithelium to be gently lifted off the mesenchyme after washing the tissue in Hanks balanced salt solution containing heat-inactivated horse serum.

smooth muscle fibres (the target cells of sympathetic axons). Our demonstration, however, that NGF mRNA is present in the epithelium of developing skin clearly indicates that the target cells of sensory neurons synthesize NGF because there are no Schwann cells in the epithelium. Although it is possible that the NGF mRNA present in mesenchyme is located in Schwann cells, there was no evidence from *in situ* hybridization that the message is restricted to bundles of nerve fibres traversing the mesenchyme.

Expression of NGF receptors

To determine the stage at which developing sensory neurons first express cell-surface receptors for NGF, 24-hour explant and dissociated cultures of E9 to E14 trigeminal neurons were incubated with ¹²⁵I-labelled NGF and processed for autoradiography to localize NGF receptors⁶. In E9 explant cultures none of the neurites was labelled (Fig. 6*a* and *b*). These neurites are the primary outgrowth from some of the earliest neurons to differentiate in the ganglion because nerve fibres

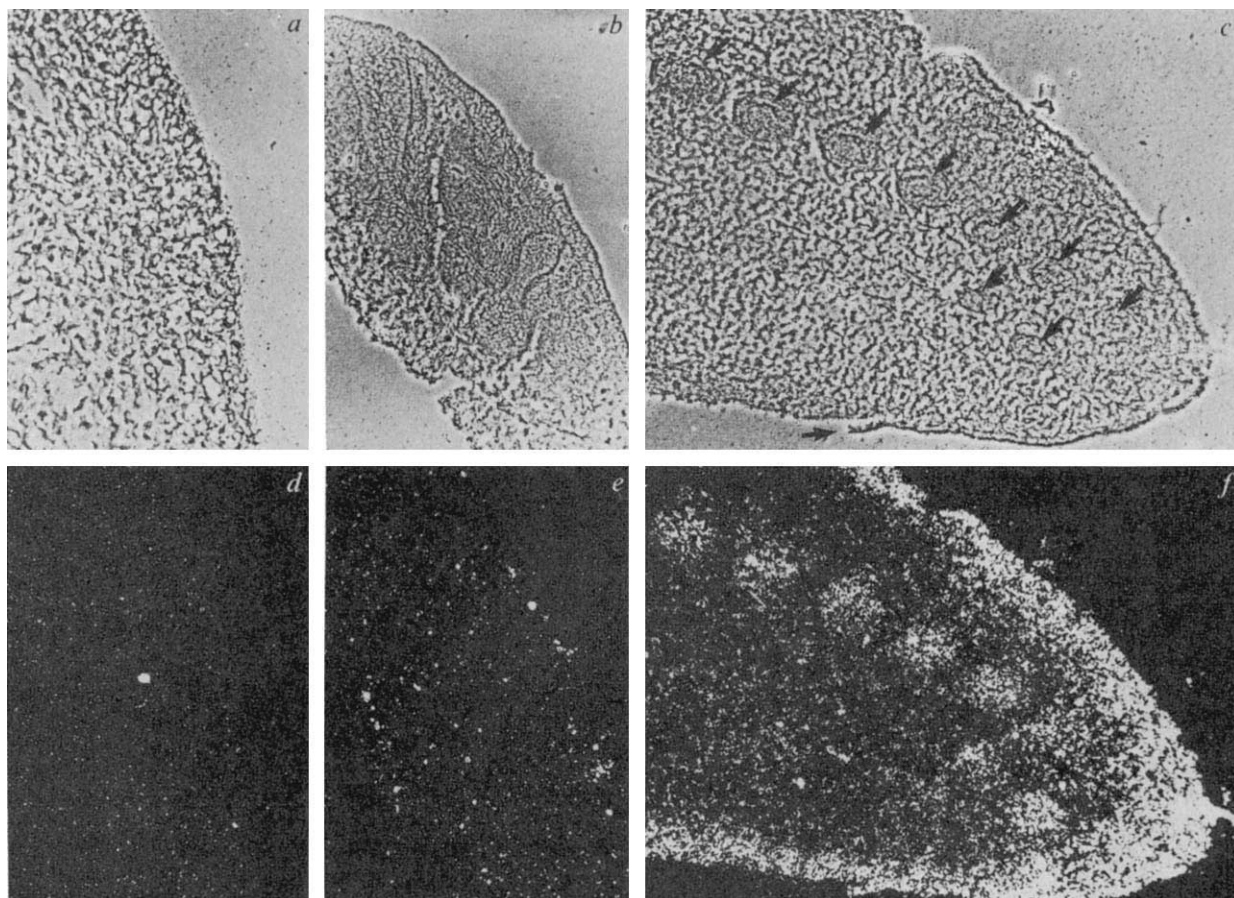


Fig. 5 Localization of NGF mRNA in tissue sections by *in situ* hybridization. Tissue was fixed, embedded, sectioned, hybridized with a 170-bp ^{35}S -labelled NGF cRNA probe and processed for autoradiography as described²⁸. *a, d*, and *b, e*, Paired phase-contrast and dark-field micrographs of autoradiographs of controls for non-specific hybridization of the probe: *a* and *d* show a section through the E10 maxillary process (in which NGF mRNA is not detectable on Northern blots) that was hybridized with the NGF cRNA probe and *b* and *e* show a section through the late E13 whisker pad (in which NGF mRNA is detectable) that was hybridized with an ^{35}S -labelled NGF RNA of similar length to the NGF cRNA probe but of the reverse polarity. In both cases the grain density over the sections is only slightly higher than background, indicating that the level of non-specific hybridization is very low. *c, f*, Section of the E13 whisker pad passing through a row of whisker follicles (small arrows) that was hybridized with the NGF cRNA probe. Both the follicular epithelium and surface epithelium, the thickness of which is illustrated by the small piece that has become detached (large arrow), are densely labelled. Although the mesenchyme shows a moderate level of labelling throughout, the region lying just beneath the surface epithelium is more heavily labelled.

have not yet emerged from the ganglion by E9 *in vivo*. Because these early fibres would not normally reach their peripheral targets *in vivo* until E10.5 (mandibular process) or E11 (maxillary process), this suggests that developing sensory neurons lack NGF receptors at the stage when their fibres are growing to their targets. This finding accords with the previous demonstration that the initiation and growth of neurites from trigeminal neurons at this early stage of development is independent of NGF²⁵. In both explant and dissociated cultures the proportion of labelled fibres and neurons increased from <30% in E10 cultures to almost 100% in E14 cultures (Fig. 6*e* and *h*). This increase in labelling is closely related to the period of development during which trigeminal nerve fibres arrive in the peripheral target-field. This relationship suggests that NGF receptors are first expressed by developing sensory neurons at the stage when their fibres reach their targets. This accords with the previous finding that sensory neurons also become responsive to NGF at this stage of development²⁵. Although the presence of a proportion of labelled neurons in E10 cultures appears to conflict with the proposal that NGF receptor expression is temporally related to target-field innervation, because in the E10 embryo trigeminal nerve fibres have not yet reached their targets, it is likely that early neurons continue to mature in culture and express NGF receptors in accordance with their actual age (E10 plus 24 hours *in vitro*) rather than the age at which they were

removed from the embryo. Direct evidence for the continued maturation of early neurons in culture is the finding that all neurons in E11 and older cultures became labelled after 48 hours of incubation (Fig. 5*h*). As many of these neurons had not made contact with their targets before culture and there was negligible neuronal death in the 24–48 hour culture period, this suggests that the expression of NGF receptors is not initiated by target encounter but is an intrinsic feature of the development of sensory neurons. In this respect there is evidence that responsiveness to NGF in chick sensory neurons appears as part of a differentiation programme several hours after the terminal mitosis of precursor cells (U. Emsberger and H.R., unpublished findings).

NGF and nerve fibre guidance

The finding that sympathetic axons grow towards an artificial source of NGF *in vivo*¹⁸ and that sensory neurites turn and grow up a concentration gradient of NGF *in vitro*^{19,20} has provided evidence for the view that NGF attracts sensory and sympathetic nerve fibres to their target-fields during development²¹. It is important to note, however, that the neurons used in all these studies had already innervated their targets *in vivo*. In this study we have found that sensory nerve fibres lack NGF receptors when they are growing to their peripheral target-field and that NGF is not present in the target-field before the arrival of the

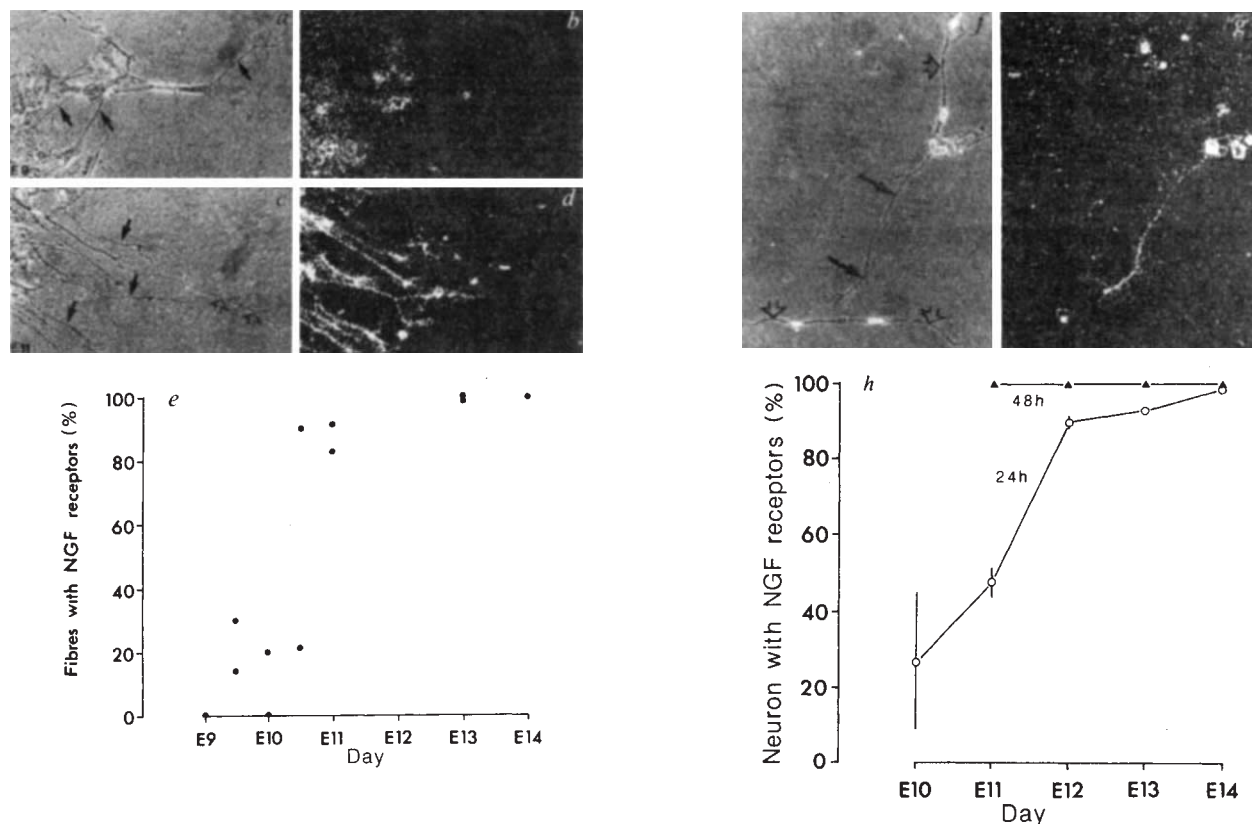


Fig. 6 Localization of NGF receptors on cultured neurons by ^{125}I -labelled NGF binding and autoradiography⁶. *a, b*, Matching phase-contrast and dark-field micrographs of part of the autoradiograph of neurite outgrowth from an E9 explant showing that none of the neurites (arrows) is labelled with silver grains. *c, d*, Phase-contrast and dark-field micrographs of neurite outgrowth from an E11 explant showing several labelled neurites (filled arrows) and an unlabelled neurite (unfilled arrows). *e*, Percentage of labelled neurites growing from explants of E9 to E14 ganglia after 24 h *in vitro*; the points indicate the mean percentage in each set of cultures. *f, g*, Phase-contrast and dark-field micrographs of an E11 dissociated culture after 24 h *in vitro* showing labelled (filled arrows) and unlabelled (unfilled arrows) neurons. *h*, Percentage of labelled neurons in dissociated cultures of E10 to E14 neurons after 24 h (○) and 48 h (▲) in culture. The mean and s.e.m. of the three or more sets of cultures established at each age are plotted and at least 100 neurons were examined in each set. No neurons were labelled in cultures that were incubated with ^{125}I -labelled NGF together with a 1,000-fold excess of unlabelled NGF.

Methods. Explant cultures and dissociated neurons were grown in F14 medium containing 10% heat-inactivated horse serum and 5% heat-inactivated fetal calf serum in 35-mm 4-well plastic Petri dishes (Greiner) pre-coated with polyornithine (0.5 mg ml⁻¹ in 0.15 M borate, pH 8.4 for 12 h) and laminin (BRL, 1 µg ml⁻¹ in phosphate buffered saline for 4 h). In contrast to trigeminal ganglion explants grown in a collagen gel^{31,32}, neurites emerged from early trigeminal ganglion explants grown on this laminin substratum in the absence of NGF. In order not to prevent the growth of neurites from NGF-dependent neurons, however, all cultures were supplemented with NGF at 5 ng ml⁻¹. After either 24 h or 48 h *in vitro*, the cultures were incubated with ^{125}I -NGF for 30 min at 4 °C, washed and processed for autoradiography as described⁶.

earliest nerve fibres at E11, some 1.5 days after these fibres started growing towards the periphery. For these two reasons we conclude that the NGF does not attract early sensory nerve fibres from sensory ganglia to their target-fields during development. It has been argued^{25,31} that the experimental demonstrations of an apparent chemotactic response of nerve fibres to NGF were due to adhesive interactions with the substratum as a result of the very high concentrations of NGF used in these experiments. Furthermore, there is evidence that early trigeminal nerve fibres are guided to their peripheral target-field by a specific attractant that is immunohistochemically distinct from NGF^{31,32}. This agent is produced by trigeminal target-field epithelium but not by the adjoining epithelium innervated by the geniculate ganglion, and unlike NGF it is synthesized when the earliest trigeminal nerve fibres are growing to the periphery.

Conclusions

Developing sensory neurons lack NGF receptors at the stage when their fibres are growing to their targets and NGF synthesis in the target-field commences with the arrival of the earliest fibres. These findings indicate that NGF does not attract sensory nerve fibres to their target-fields during development, but is involved in the target-controlled survival of neurons. This contention is supported by our finding that the concentration of

NGF mRNA in the target-field correlates with the density of innervation; it is highest in the epithelium, lower in the subjacent mesenchyme and lowest in the deep mesenchyme.

Our findings also raise the question of how the synthesis of NGF and NGF receptors is regulated during development. The onset of both NGF and NGF receptor synthesis is closely related to the start of target-field innervation. Whereas the function of neuron-target interaction in the onset of NGF synthesis remains unknown, receptor expression is not triggered by target encounter, rather it seems to occur as part of the intrinsic differentiation programme of the neuron.

We thank Yves Barde for helpful discussion of this work.

Received 20 November 1986; accepted 28 January 1987.

1. Oppenheim, R. W. in *Studies in Developmental Biology* (ed. Cowan, W. M.) 74-133 (Oxford University Press, Oxford, 1981).
2. Hamburger, V. & Levi-Montalcini, R. *J. exp. Zool.* **162**, 133-160 (1949).
3. Levi-Montalcini, R. & Angeletti, P. U. *Physiol. Rev.* **48**, 534-569 (1968).
4. Thoenen, H. & Barde, Y.-A. *Physiol. Rev.* **60**, 1284-1335 (1980).
5. Davies, A. M. & Lindsay, R. M. *Dev. Biol.* **111**, 62-72 (1985).
6. Rohrer, H. & Barde, Y.-A. *Dev. Biol.* **89**, 309-315 (1982).
7. Hendry, I. A. & Campbell, J. J. *Neurocytol.* **5**, 351-360 (1976).
8. Hamburger, V., Brunso-Bectold, J. K. & Yip, J. W. *J. Neurosci.* **1**, 60-71 (1981).
9. Gorin, P. D. & Johnston, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5382-5386 (1979).
10. Heumann, R., Korsching, S., Scott, J. & Thoenen, H. *EMBO J.* **3**, 3183-3189 (1984).
11. Shelton, D. L. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7951-7955 (1984).
12. Hendry, I. A., Stoeckel, K., Thoenen, H. & Iversen, L. L. *Brain Res.* **68**, 103-121 (1974).

13. Stoeckel, K., Schwab, M. & Thoenen, H. *Brain Res.* **89**, 1-14 (1974).
14. Angeletti, R. H. & Bradshaw, R. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2417-2420 (1971).
15. Scott, J. *et al. Nature* **302**, 538-540 (1983).
16. Ullrich, A., Gray, A., Berman, C. & Dull, T. J. *Nature* **303**, 821-825 (1983).
17. Darling, T. L. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **48**, 427-434 (1983).
18. Menescini-Chen, M. G., Chen, J. S. & Levi-Montalcini, R. *Archs ital. Biol.* **116**, 53-84 (1978).
19. Letourneau, P. C. *Dev Biol.* **66**, 183-196 (1978).
20. Gundersen, R. W. & Barrett, J. N. *Science* **206**, 1079-1080 (1979).
21. Levi-Montalcini, R. *Prog. Brain Res.* **45**, 235-258 (1976).
22. Renehan, W. E. & Munger, B. L. *J. comp. Neurol.* **246**, 129-145 (1986).
23. Iggo, A. *Br. med. Bull.* **33**, 97-102 (1977).
24. Iggo, A. & Andres, K. H. *A. Rev. Neurosci.* **5**, 1-31 (1982).
25. Davies, A. M. & Lumsden, A. G. S. *J. comp. Neurol.* **223**, 124-137 (1984).
26. Davies, A. M. & Lumsden, A. G. S. *J. comp. Neurol.* **253**, 13-24 (1986).
27. Korsching, S. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3513-3516 (1983).
28. Bandtlow, C., Heumann, R., Schwab, M. & Thoenen, H. *EMBO J.* (in the press).
29. Rush, R. A. *Nature* **312**, 364-367 (1984).
30. Finn, P. J., Ferguson, I. A., Renton, F. J. & Rush, R. A. *J. Neurocytol.* **15**, 169-176 (1986).
31. Lumsden, A. G. S. & Davies, A. M. *Nature* **306**, 786-788 (1983).
32. Lumsden, A. G. S. & Davies, A. M. *Nature* **323**, 538-539 (1986).
33. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

Three-dimensional structure of a complex of antibody with influenza virus neuraminidase

P. M. Colman, W. G. Laver*, J. N. Varghese, A. T. Baker, P. A. Tulloch, G. M. Air† & R. G. Webster‡

CSIRO Division of Protein Chemistry, 343 Royal Parade, Parkville, 3052, Australia

* John Curtin School of Medical Research, Australian National University, Canberra 2601, Australia

† Department of Microbiology, University of Alabama, Birmingham, Alabama 35294, USA

‡ St Jude Children's Research Hospital, Memphis, Tennessee 38101, USA

The structure of a complex between influenza virus neuraminidase and an antibody displays features inconsistent with the inflexible 'lock and key' model of antigen-antibody binding. The structure of the antigen changes on binding, and that of the antibody may also change; the interaction therefore has some of the character of a handshake.

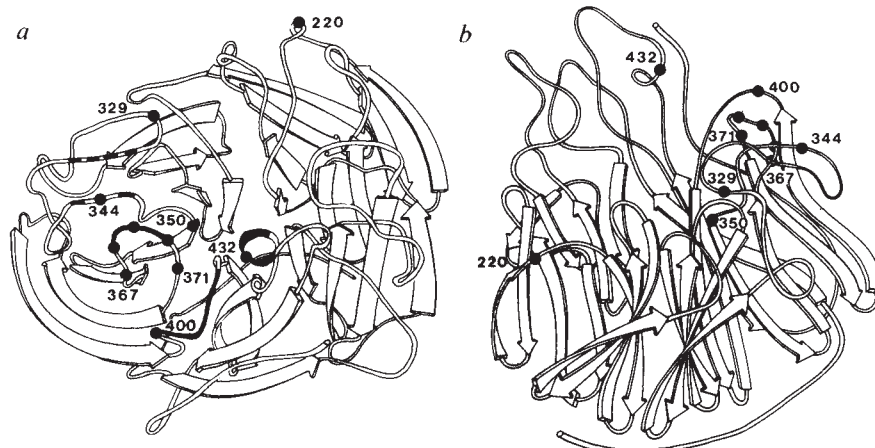
WE report here the first analysis by X-ray diffraction of the three-dimensional structure of a viral antigen complexed with an antibody Fab fragment. The antigen is the influenza virus enzyme neuraminidase (reviewed in ref. 1).

Sufficient data on the chemical² and spatial³⁻⁶ structure of antibodies emerged in the 1970s to provide a basis for understanding how variation in antibody structure occurs, but the question of how different antibodies accommodate different macromolecular epitopes remains unresolved. The antigen binding fragments (Fab) of immunoglobulins consist of a light chain (L) and the N-terminal half of the heavy (H) chain. On each chain there are two globular domains of ~100 amino acids, the N-terminal domain in each chain being variable (V_L and V_H) and the C-terminal domain conserved (C_L and C_H1) in their amino-acid sequences. V_L and V_H domains are associated with each other, forming a variable module and C_L and C_H1 form the constant module. Three complementarity-determining regions (CDR1, 2 and 3) from each of the V_L and V_H domains are clustered at the extremity of the Fab arms, and

they determine the binding specificity of an immunoglobulin molecule.

The three-dimensional structures of five free Fab fragments⁷⁻¹¹ and one in complex with lysozyme¹² have been reported. Four of the structures, New⁷, Kol⁸, McPC603⁹ and J539¹⁰, are well refined and form the basis of the following generalizations which also appear valid for the other two, S10/1 (ref. 11) and D1.3 (ref. 12). The pairing of V_L and V_H domains is determined by amino acids from both the framework region and the CDRs. Most of the surface area buried in the interface derives from conserved residues and this observation explains the largely conserved geometry of V_L-V_H pairing¹³. Small differences in pairing presumably result from interactions among the CDR amino acids¹⁴. Analysis of the V_L-V_H contact surface shows it to be unlike other interfaces between β-sheets¹⁵. Amino acids from the outermost strands of the two sheets fold into the interface where they contribute the bulk of the buried surface. C_L-C_H1 dimers are similarly conserved in their association pattern¹⁶. More variable is the so-called elbow angle^{17,18}

Fig. 1 Schematic diagrams of chain fold in N2 and N9 neuraminidase viewed down the 4-fold axis (a) and perpendicular to this axis (b). The symmetry axis is bottom right in a and standing vertical at the left rear in b. The view of the neuraminidase in b is similar to that shown in Fig. 5. Tagged residues and adjacent chain segments are referred to in the text. N2 numbering is used. In a, the side chains of amino acids 368-370 point towards the viewer, whilst that of Arg 371 points away and into the catalytic site located above and to the right of C_α 371. Mutations at positions 367, 369, 370, 400, and 432, as detailed in Fig. 2, abolish the binding of NC41 antibody to neuraminidase, whereas mutations at 368 and 329 reduce that binding. A mutation at 220, which falls outside the NC41 epitope, has no effect on NC41 binding to neuraminidase (see Fig. 2). In a, solid chain segments show regions in contact with the NC41 antibody; contact assignments in the broken solid segments are tentative.



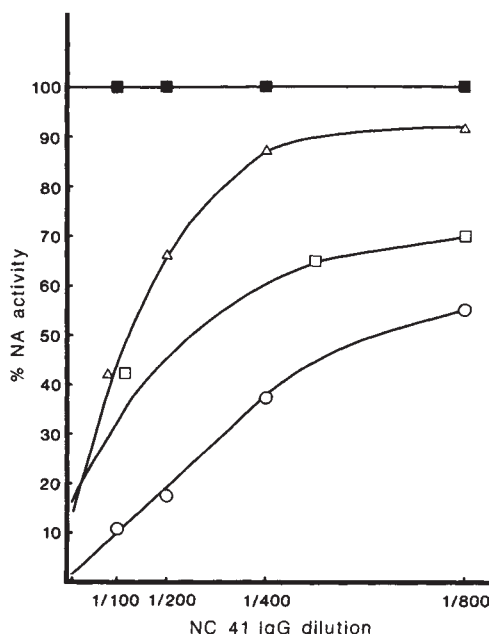


Fig. 2 Neuraminidase inhibition (NI) curves of monoclonal antibody NC41 acting on N9 neuraminidase (NA) and a number of variants selected with different monoclonal antibodies to N9 neuraminidase. Assays used fetuin as a substrate, as described⁵⁸. The amino-acid sequence changes of the variants are: ■, 434 (432 in N2 numbering) K → N; 400 (400) N → K; 371 (370) S → L; 370 (369) A → D; 368 (367) S → N; □, 369 (368) I → R; △, 331 (329) N → D (designated as OX2 in text); ○, 222 (220) R → Q (this curve is identical to wild-type). The amino-acid sequence changes were determined by sequencing cDNA of the viral RNA strand encoding the neuraminidase as described⁵⁹. The changes are shown on the 3-D structure of the monomer in Fig. 1.

describing the angle between the local axes of symmetry in the V and C modules.

The identification of antigenic regions on protein molecules¹⁹ has been attempted by a variety of methods^{20–23} but the structural basis of antigenicity remains unclear. Current models of antigen-antibody association are based on classical ideas of lock and key complementarity. But it has been suggested²⁴ that the antigenic regions of proteins are located in the more flexible chain segments, which may adopt a configuration that allows antibody binding. It is not known whether binding to antigen has any effect on the structure of an antibody²⁵, nor whether, if such an effect exists, it has biological significance. For the Fab-lysozyme complex¹², the parts of the lysozyme that make contact with the antibody are not the more flexible regions of free lysozyme. No deformation of the antigen nor any structural change in the Fab was observed¹².

The antigen in this study is the neuraminidase of influenza virus. It is a tetramer of subunits with relative molecular mass 60,000 (M_r 60 K) with circular symmetry, and is attached by a stalk to the viral membrane. Neuraminidase 'heads' can be liberated from the virus by proteolysis²⁶. There is antigenic variation in the neuraminidase between different strains of virus. Antibodies against neuraminidase do not neutralise virus infectivity, but do modify the disease in favour of the host²⁷. The three-dimensional structures of neuraminidase heads of subtypes²⁸ N2 (refs 29, 30) and N9 (A.T.B., J.N.V., W.G.L., G.M.A. and P.M.C., manuscript in preparation) are known and a schematic of the chain fold is shown in Fig. 1. Neuraminidase of subtype N9 from an avian influenza virus, G70c³¹, has been used in this study. Its structure is very similar to the N2 enzyme as expected from the sequence homology of 50% (ref. 32).

We previously reported electron and low resolution X-ray diffraction studies of Fab-neuraminidase complexes³³ and, more

recently, the crystallization of another complex (with NC41 Fab) which diffracts X-rays to beyond 3 Å resolution³⁴. The NC41 antibody can suppress the yield of virus from infected cells and therefore allows the selection of antigenic variants. Neuraminidase-inhibition curves (Fig. 2) show that some variants with single sequence changes are not inhibited at all by NC41 antibody whereas others are partially inhibited. The location of these sequence changes on the three-dimensional structure of neuraminidase is shown in Fig. 1.

Protomer structure

Electron microscopy of negatively stained protomers of complex show tetrameric neuraminidase heads with four Fab fragments attached (data not shown). Their image appearance is very similar to that of protomers of N9 neuraminidase complexed with the Fab fragments of antibodies 32/3 and NC35 (ref. 33) and may be described as a square box ($100 \times 100 \times 60$ Å) with four antennae (Fab molecules) attached to one surface on the outer corners. The overall shape of the protomer from the X-ray diffraction study (Fig. 3) is consistent with the electron microscope images. The Fab arms subtend an angle of 45° to the plane of the tetrameric neuraminidase head, slightly foreshortened in the electron microscope image³³.

Surface loops of the neuraminidase in contact with the CDRs are 368–370, 400–403, 430–434 and parts of 325–350. Of these, the conformation of the 325–350 loop is still tentative. It was the most difficult region of the original N2 neuraminidase structure²⁹ to interpret and remains ambiguous in the N9 neuraminidase structure (A.T.B. *et al.*, manuscript in preparation). Such instances of ill-defined structure are typical of surface loops of polypeptides subject to thermal motion or statistical disorder.

Early indications on the distribution of temperature factors in the uncomplexed N2 neuraminidase, based on crystallographic refinement of that structure at 2.2 Å resolution and an R -factor of 0.293, show above-average thermal parameters in chain segments 109–111, 141–142, 221–222, 243–247, 305–309, 316–319, 327–331, 338–343, 365–369, 399–402, 429–437 and 461–469 (J.N.V. and P.M.C., unpublished data). Elevated B values around residue 330 may correlate with inaccurate modelling as discussed above. All the segments listed here are surface loops. Note that only one of the hotter loops is underneath the neuraminidase head. We emphasize that the temperature factor analysis is preliminary, and indeed relates to N2, not N9, neuraminidase.

The monoclonal variant of N9 neuraminidase known as OX2 (Asn 329 → Asp 329, N2 numbering) crystallizes isomorphously with wild-type N9-NC41 Fab complex³⁴. NC41 inhibits the OX2 neuraminidase variant less than it does wild-type (Fig. 2). Refinement of the N9-NC41 Fab structure using data collected from the OX2-NC41 Fab complex yields an R -factor of 0.36 for a model with tight geometry. This result is comparable with the current R -factor against the native dataset and confirms our earlier conclusion³⁴ that the binding of NC41 Fab to wild-type and OX2 N9 neuraminidases is isosteric. Either some bonding 'glue' is absent or repulsive forces have been introduced by the mutation (Fig. 2), but there is no detectable rearrangement of the antibody on the antigen at this stage. The precise location of residue 329 (N2 numbering) in the N2 and N9 structures is unclear. Refinement of the OX2-NC41 Fab structure should eventually reveal the structural basis for the reduced binding of NC41 to OX2 compared with wild-type N9.

The third CDRs on both heavy and light chains are not yet completely modelled, and the conformation of the 325–350 region on the antigen remains uncertain, so we cannot identify the amino acids that interact in the interface. All the CDRs, with the possible exception of light-chain CDR1 appear to make contact with the epitope. A preliminary assessment is that there are no more contacting residues than the 16 or 17 observed in the Fab-lysozyme structure¹².

Quaternary structure of the Fab

The amino-acid sequence of the Fab is known (G.M.A., unpublished data). The assignment of heavy and light chains in the Fab structure is consistent with side-chain densities at homologous positions on the two chains, such as heavy-chain (H) Trp 47 and light-chain (L) Leu 46, H Val 37 and L Tyr 36, and, H Trp 103 and L Phe 98. Heavy-atom binding sites also show an asymmetry which is consistent with these observations.

The quaternary structures of known Fab fragments were compared as described in Table 1. The lower left triangle of entries indicates that C_L - C_{H1} pairing is a function of immunoglobulin chain subtype since γ_1 : λ chains (as in Kol and New) show a consistent mode of association as do α or γ_{2a} : κ chains (as in M603 and NC41). Small but significant alterations in V-domain pairing are observed between Kol, New and M603 (Table 1, upper right). This variation in V_L - V_H association has been attributed to the fact that the CDRs contribute partially to the dimer interface between the two domains in the V-module¹⁴, although in the case of Kol we suggest that intimate crystal contacts involving the CDRs may contribute.

Comparison of these three Fabs with NC41 complexed with neuraminidase shows, in each case, a significantly larger difference in V_L - V_H association. Compared with Kol, the distance between H CDR2 and L CDR3 is larger and the relative positions of the CDRs around the antigen-binding surface are altered by up to 4 Å. A movement of this magnitude is of the order of the distance between adjacent α -carbon atoms on a polypeptide and may be considered as dramatic as a sequence insertion or deletion in a structurally critical part of the molecule. As well as the bulk movement of domains described here, the CDR loops may be flexible and able to adapt further to an epitope.

Building an atomic model into the penultimate electron density map (that in which no phase information from the Fab structure had been included) required the rotation of H Trp 47 180° around C_α - C_β compared with its position in Kol. M603 and New are the same as Kol at this point. The positions of the loops at residue 40 on both heavy and light chains also required remodelling. Therefore the reorganization of the interface through sliding of domains has consequences for the domain structure of the V_L and V_H .

The sequence of the NC41 V_L and V_H domains shows that those framework amino acids on both the light chain (Tyr 36, Gln 38, Pro 44, Tyr 87, Phe 98) and the heavy chain (Val 37, Gln 39, Leu 45, Trp 47, Phe 91, Ala 93, Trp 103) that are buried in the interface¹⁵ are conserved as expected. There is nothing in the primary sequences of the CDRs of NC41 to suggest a cause of the unusual domain association. An independent analysis of the NC41 sequence confirms this view (C. Chothia, personal communication).

Two other Fab structures are known to fit the normal pattern of V_L - V_H association, the free antineuraminidase monoclonal antibody S10/1 (ref. 11) and the complexed anti-lysozyme monoclonal antibody D1.3 (ref. 12). Among light-chain dimers the picture is less clear. Although several examples of V_L - V_H -like association have been observed^{3,35-37}, there are two clear exceptions³⁸⁻³⁹, one of which is said to be caused by the CDR residue L 91, or, in our view, perhaps by a His at residue 38.

If changes in the amino-acid sequence of the CDRs (which are only a small part of the interface) can modulate the pairing pattern, binding of the six CDRs to a macromolecular surface might similarly perturb the V_L - V_H interface. Without knowing the three-dimensional structure of the free NC41 Fab we cannot definitely conclude that the antibody structure has changed on binding antigen, but we believe that this explanation is more plausible than the alternative, which requires that special sequences in the CDRs alone have determined the pairing pattern.

Other data support conformational changes in antibodies when antigen or hapten binds⁴⁰⁻⁴⁴. Changes in circular dichro-

Table 1 Quaternary structure comparisons of Fabs

	NEW	M603	KOL	NC41
NEW		4.2 (0.7)	6.5 (0.4)	12.2 (1.0)
M603	12.7 (2.3)		5.0 (0.3)	8.8 (1.3)
KOL	2.5 (0.0)	12.4 (2.4)		12.1 (0.7)
NC41	11.5 (0.4)	2.1 (0.7)	10.8 (2.1)	

Comparison of four different Fab quaternary structures. Upper right triangle: V_L domain of the row Fab was mapped into the V_L domain of the column Fab using the C_α coordinates of residues 33-39, 43-47, 84-90, 98-104 (Kabat² numbering). These amino acids were chosen because they define the interface between V_L and V_H domains¹⁵. The calculated transformation was applied to the V_H domain of the row Fab and the additional rotation in degrees (and translation in Å) to optimize overlap of the two V_H domains in question was calculated and is given above. C_α coordinates used for overlap of V_H domains were 34-40, 44-48, 88-94 and 103-109. Lower left triangle: C_L domain of the column Fab was mapped into the C_L domain of the row Fab using the C_α coordinates of residues 118-124, 131-137, 160-162 and 174-178 (Kabat² numbering). These amino acids were chosen because they define the interface between C_L and C_{H1} domains¹⁶. The calculated transformation was applied to the C_{H1} domain of the column Fab and the additional rotation in degrees (and translation in Å) to optimize overlap of the two C_{H1} domains are shown. C_α coordinates used for overlap of C_{H1} domains were 121-125, 139-145, 172-179 and 186-190. Amino acids from the interface were used because if this set is expanded to include other parts of the domain structure, the correlation between small rotation angles and similar chain classes that is observed for the C module is less good. The largest r.m.s. distance between domain structure alignments is 0.7 Å. A similar comparison was done for the V modules of the M603 (antiphosphocholine) and uncomplexed S10/1 (antineuraminidase) Fab fragments. No measurable difference in the way the V_L and V_H domains associate in those antibodies was detected. A comparison elsewhere of Kol and New V modules⁸ yields a value of 9°. The value reported here (6.5°) is smaller because it is uninfluenced by structural differences outside of the interface zone.

ism⁴⁰ or circular polarization of fluorescence⁴² indicate an altered environment for aromatic residues after antigen binds to antibody, and kinetic data support a bi-molecular process with distinct conformational states for bound and free Fab^{41,43,44}. Both of these aspects are embodied in the sliding V_L - V_H model presented here.

Direct structural evidence has been presented for hapten-induced structural changes in a Bence-Jones protein^{45,46}. In that case hydrophobic ligands can apparently penetrate the V_L - V_L interface and signal their presence to the C-module. Strain introduced in the structure of the C-module can be relaxed by reducing and re-oxidizing the disulphide bond between the C-terminal regions of the two light chains⁴⁶. Table 1 indicates that in the neuraminidase-NC41 complex no significant changes in the quaternary structure of the C-module have occurred.

Analysis of the pseudo-symmetry axes of the V and C modules shows that the angle between these axes, the elbow angle, is ~150°, which is intermediate in the observed range of 130-180° for other Fab structures. The sense of the bending is the same as in all other Fab fragments, with V_H and C_{H1} domains closer together than the V_L and C_L domains.

Conformational changes in the antigen

The neuraminidase upper surface loop involving residues 367-371 was remodelled to better fit the density as observed in the complex. Adjacent parts of the loop which were also omitted from the phasing process (residues 364-366 and 372-373) fit the density as if they are rigidly attached to the N9 core. The C_α atoms of residues 368-371 have been moved by 1 Å or more from their position in N9 neuraminidase (Fig. 4). Least-squares refinement of the structure to $R=0.35$, gave a mean shift of 0.4 Å in all C_α atoms with a standard deviation of 0.2 Å. The current position of the 370 loop (Fig. 4) shows displacements of more than 1.2 Å for C_α atoms 369-371 compared with their

Fig. 3 Stereo images of the C_α skeleton of the protomer, showing four Fab molecules attached to one tetrameric neuraminidase 'head'. Neuraminidase, purple; heavy chain, green; light chain, blue. View with the protomer 4-fold axis vertical, but tipped towards the viewer. The crystals of N9-NC41 Fab belong to the space group P4₂2 with $a = 167$ Å, $c = 124$ Å (ref. 34).

There is one quarter of a protomer per asymmetric unit and 70% of the cell volume is occupied by solvent. Data were collected photographically to 3 Å resolution. In all, 67,976 measurements of 25,241 independent reflections, representing 76% of the data in the 3 Å sphere, were merged with an R -factor on intensities of 0.12. Phasing was initiated through two heavy atom derivatives (potassium tetra-chloro-platinate and diamino-dinitro-platinum). An envelope was computed⁶⁰ on the basis of a 5 Å resolution electron-density map allowing the location of the neuraminidase tetramer on the crystallographic 4-fold axis to be determined first by inspection and subsequently by correlation methods in real space and rigid body least-squares refinement⁶¹, resulting in a correlation coefficient of 0.204 between F_{obs} and F_{calc} for data in the range 5–3.5 Å. It was now clear which surface regions of the neuraminidase were in contact with the remaining electron density in the image, and, in subsequent use of neuraminidase as partial structure in the phasing process, those regions were excised. A second electron density map, now at 3.5 Å resolution, was phased on the two derivatives (one to 5 Å and one to 3.5 Å), the neuraminidase partial structure, and solvent flattening from a redetermined envelope. The four domains of the Fab fragment were recognised in this map and fitted independently by real-space correlation methods with the known Fab structure Kol⁸. Rigid-body least-squares refinement with five groups, neuraminidase, V_H , C_{H1} , V_L and C_L produced shifts of up to 1.5° in group orientations and led to an R -factor of 0.425 and a correlation coefficient of 0.308 for data between 5 and 4 Å resolution. The structure is based on a 3 Å electron density map, which is phased in addition to the information listed above with the structure of the Fab fragment excluding the complementarity-determining loops. Thus there is no structural prejudice in the image of those parts of the antigen or antibody that are near, or in, the bonding interface. Further refinement⁶² has since reduced the residual to 0.35 for data between 6 and 3 Å resolution. The r.m.s. error in the atomic positions is of the order of 0.5 Å at this stage.

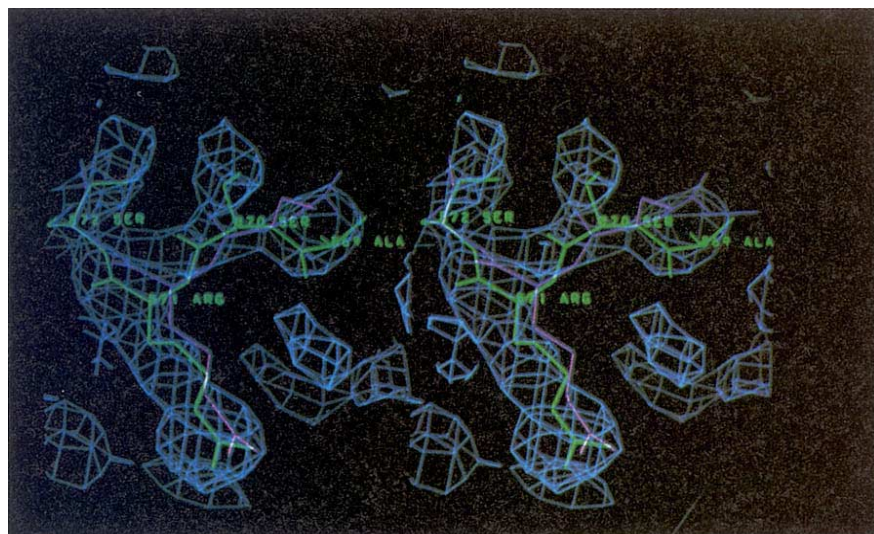
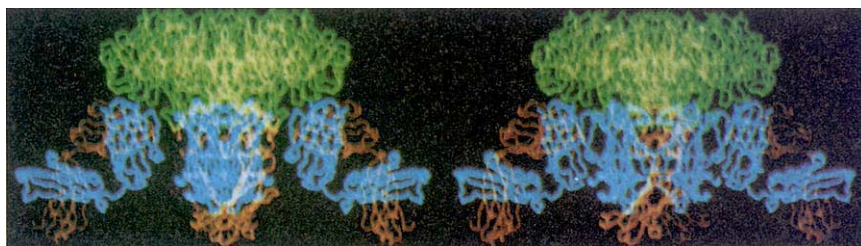
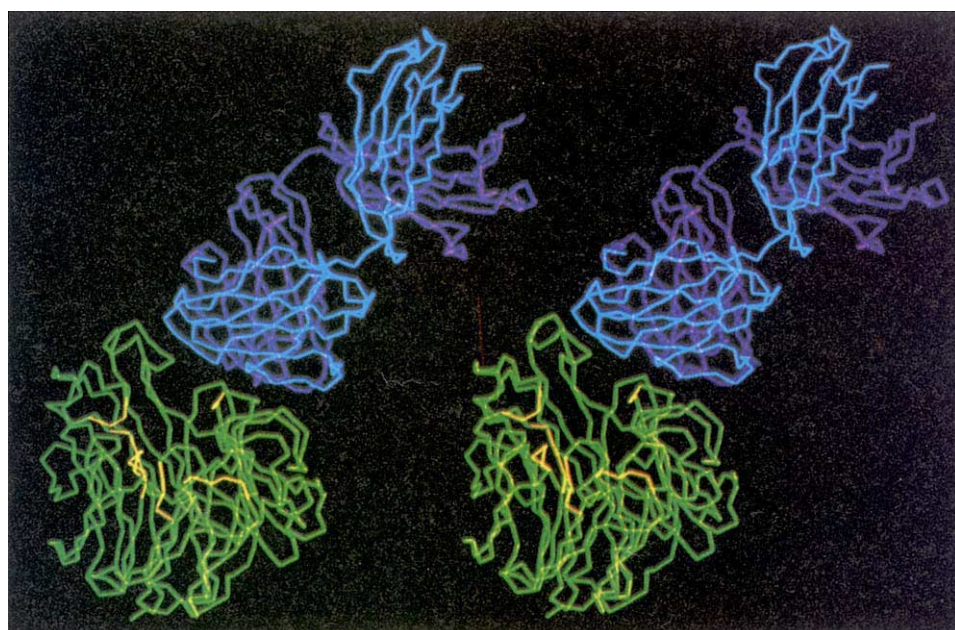


Fig. 4 Stereo image of the electron density map (2Fb-Fc, $R = 0.35$) of the neuraminidase loop 368–372 in the complex with NC41 Fab. Residues here are labelled 868–872. The yellow model shows the current fit of the loop in the complex structure. The purple model is the position of this loop in uncomplexed N9 neuraminidase.

Fig. 5 Stereo image of one quarter of the protomer, showing the Fab bound to the side of the active site cavity on the enzyme. The enzyme active centre is facing the viewer, below and to the left of the antibody binding site, and is coloured yellow. The neuraminidase perspective is as in Fig. 1b. Neuraminidase, green; heavy chain, purple; light chain, blue.



position in free N9 neuraminidase. Thus the antigen is distorted by interaction with the antibody.

Most monoclonal anti-neuraminidase antibodies inhibit neuraminidase activity against fetuin, but only some of these also inhibit activity against the trisaccharide sialyl-lactose⁴⁷. Although NC41 Fab cannot sterically block entry of a trisaccharide into the active site (Fig. 5), NC41 does inhibit this activity (R.G.W., manuscript in preparation). We suggest that one possible mechanism for enzyme inactivation is the small displacement of Arg 371, which points directly into the catalytic site. We cannot yet rule out other interpretations, such as the effect of the antibody on steering substrates and products into or out of the active site. It has been shown that changing Arg 371 for Lys by site-specific mutagenesis results in the loss of 90–95% of enzyme activity, although the mutated protein is correctly folded (M. R. Lentz, R.G.W. and G.M.A., manuscript in preparation).

Deformation of antigens by antibodies has been proposed to explain a variety of experimental data, including single-hit kinetics for neutralization of polio virus⁴⁸ and altered binding affinity for second antibodies binding noncompetitively to the antigen⁴⁹. The capacity of anti-apomyoglobin antibodies to, effectively, expel the haem from myoglobin has long been known⁵⁰. There are also many reports of antibodies inhibiting enzymes⁵¹. Our results directly demonstrate that conformational changes can be induced in antigens by antibodies. Furthermore, the change observed here, although small, correlates with the inactivation of the neuraminidase towards sialyl-lactose. Other anti-neuraminidase antibodies that inhibit neuraminidase activity only towards large substrates and, like NC41, can select antigenic variants, probably inhibit enzyme activity by blocking access of large substrates to the active site. We are testing the hypothesis that such antibodies do not distort the active site of the enzyme by an X-ray diffraction study of the N9-NC10 Fab complex³⁴. Unlike NC41, this antibody does not inhibit enzyme activity on sialyl-lactose. Antigen distortion is likely to be but one of a number of possible mechanisms for virus neutralization⁵².

Conclusions

The structure of a complex between an antibody and influenza virus neuraminidase shows features inconsistent with a rigid 'lock and key' model for antibody-antigen interactions. In contrast to the findings from a lysozyme-antibody complex¹² we observe (1) an unusual V_L - V_H pairing in the V module of the Fab, and (2) local perturbation of the antigen at the centre of the epitope. The interaction therefore has some of the character of a handshake. NC41 Fab has not yet been crystallized, and we can only suppose on the basis of sequence data that its structure is similar to the common pattern among all six other Fab structures in the literature. We propose that, during antigen binding, the V_L and V_H domains of the antibody slide at their interface. The extent of sliding is small, but sufficient to move the CDRs by at least 3 Å. If such a structural transition has occurred, it is presumably of low energy and therefore low cost to the binding affinity for the antigen. The observation that V_L - V_H dimers associate in a way not found in associations of other β -sheet structures has led to the suggestion that this novel association might produce static or dynamic properties important for antigen binding¹³. The ability of domains to slide could expand the range of specificities of a single antibody species. Cross-reactivity of monoclonal antibodies with proteins unrelated to the immunizing antigen⁵³ could occur by different extent, or direction, of domain sliding.

Diversity of antibody specificities is generated by libraries of germline gene sequences, by somatic events and by combination of different H and L chains⁵⁴. If the V module can generate diversity through V_L - V_H sliding, this presents yet another mechanism for fine-tuning the specificity of an antibody. We do not suggest that all antigen-antibody interactions will involve

reorganization of the V_L - V_H interface. The lysozyme-Fab complex¹² apparently does not. Thus the only biological significance we attach to V_L - V_H sliding is that, in some instances, it may facilitate the formation of a functional immune complex.

A necessary, but apparently insufficient, condition to induce sliding is the interaction of the epitope with CDRs from both H and L chains. It may be that the larger radius of curvature of neuraminidase, compared with lysozyme, restricts the capacity of the Fab to bind in its ground state.

Sequence data on the T-cell receptor are consistent with the α - and β -chains being immunoglobulin-like^{55,56}. We note further that those sequences contain the characteristics of V_L and V_H sequences that have been attributed to the novel association of V_L and V_H domains¹⁵. If the principles of recognition of antigens by T- and B-cell receptors are indeed similar⁵⁷, then induced fit of antigen to receptor might also be a feature of cellular immunity.

As the two reported antigen-antibody complex structures are rather different, more data on similar systems are required to establish the importance of V_L - V_H sliding to immune recognition.

We thank Paul Davis for computing assistance, Bert van Donkelaar and Janet Newman for technical assistance, Cyrus Chothia for communicating results in advance of publication and the Australian Overseas Telecommunications Commission for donating international dialling facilities. Some of the OX2-NC41 data were collected at SSRL and we thank Paul Phizackerly for assistance. This work was supported by the US National Institutes of Health.

Note added in proof: The NC41 V_H belongs to the recently characterized V_H family IX (Winter, E., Radbuch, A. and Krawinkel, V. *EMBO J.* 4, 2861–2867 (1985)).

Received 11 November 1986; accepted 3 February 1987.

- Colman, P. M. & Ward, C. W. *Curr. Top. Microbiol. Immun.* 114, 178–255 (1985).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Proteins of Immunological Interest* (US Department of Health and Public Services, Washington DC, 1983).
- Schiffer, M., Girling, R. L., Ely, K. R. & Edmundson, A. B. *Biochemistry* 12, 4620–4631 (1973).
- Poljak, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* 70, 3305–3310 (1973).
- Segal, D. M. *et al. Proc. natn. Acad. Sci. U.S.A.* 71, 4298–4302 (1974).
- Matsushima, M. *et al. J. molec. Biol.* 121, 441–459 (1978).
- Saul, F., Amzel, L. M. & Poljak, R. J. *J. biol. Chem.* 253, 585–597 (1978).
- Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. *J. molec. Biol.* 141, 369–392 (1980).
- Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. *J. molec. Biol.* 190, 593–604 (1986).
- Suh, S. W. *et al. Proteins: Structure, Function and Genetics*, 1, 74–80 (1986).
- Colman, P. M. & Webster, R. G. in *Biological Organisation: Macromolecular Interactions at High Resolution* (ed. Burnett, R.) 125–133 (Academic, New York, 1987).
- Amit, A. G., Marriuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Science* 233, 747–753 (1986).
- Novotny, J. & Haber, E. *Proc. natn. Acad. Sci. U.S.A.* 82, 4592–4596 (1985).
- Davies, D. R., Padlan, E. A. & Segal, D. M. *A. Rev. Biochem.* 44, 639–667 (1975).
- Chothia, C., Novotny, J., Brucoleri, R. & Karpus, M. *J. molec. Biol.* 186, 651–663 (1986).
- Padlan, E. A., Cohen, G. H. & Davies, D. R. *Molec. Immun.* 23, 951–960 (1986).
- Colman, P. M., Deisenhofer, J., Huber, R. & Palm, W. *J. molec. Biol.* 100, 257–282 (1976).
- Davies, D. R. & Metzger, H. *A. Rev. Immun.* 1, 87–117 (1983).
- Benjamin, D. C. *et al. A. Rev. Immun.* 2, 67–101 (1984).
- Crompton, M. J. & Wilkinson, J. M. *Biochem. J.* 94, 545–556 (1965).
- Smith-Gill, S. J. *et al. J. Immun.* 128, 314–322 (1982).
- Lerner, R. A. *Nature* 299, 592–596 (1982).
- Tainer, J. A. *et al. Nature* 312, 127–134 (1984).
- Westhof, E. *et al. Nature* 311, 123–126 (1984).
- Metzger, H. *Contemp. Top. molec. Immun.* 7, 119–152 (1978).
- Laver, W. G. *Virology* 86, 78–87 (1978).
- Schulman, J. L. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 373–393 (Academic, New York, 1975).
- Webster, R. G., Laver, W. G. & Air, G. M. in *Genetics of Influenza Viruses* (ed. Palese, P. & Kingsbury, B. W.) 127–168 (Springer, New York, 1983).
- Varghese, J. N., Laver, W. G. & Colman, P. M. *Nature* 303, 35–40 (1983).
- Colman, P. M., Varghese, J. N. & Laver, W. G. *Nature* 303, 41–44 (1983).
- Laver, W. G., Colman, P. M., Webster, R. G., Hinshaw, V. S. & Air, G. M. *Virology* 137, 314–323 (1984).
- Air, G. M., Ritchie, L. R., Laver, W. G. & Colman, P. M. *Virology* 145, 117–122 (1985).
- Tulloch, P. A. *et al. J. molec. Biol.* 190, 215–225 (1986).
- Laver, W. G., Webster, R. G. & Colman, P. M. *Virology* 156, 181–184 (1987).
- Epp, O. *et al. Biochemistry* 14, 4943–4952 (1975).
- Epp, O. *et al. Eur. J. Biochem.* 45, 513–524 (1974).
- Colman, P. M., Schramm, H. J. & Guss, J. M. *J. molec. Biol.* 116, 73–79 (1977).
- Furey, W., Wang, B. C., Yoo, C. S. & Sax, M. *J. molec. Biol.* 167, 661–692 (1983).
- Chang, C. H. *et al. Biochemistry* 24, 4890–4897 (1985).
- Holowka, D. A., Strosberg, A. D., Kimball, J. W., Haber, E. & Cathou, R. E. *Proc. natn. Acad. Sci. U.S.A.* 69, 3399–3403 (1972).
- Levison, S. A., Hicks, A. N., Portman, A. J. & Dandliker, W. B. *Biochemistry* 14, 3778–3786 (1975).

42. Schlessinger, J., Steinberg, I. Z., Givol, I. D., Hochman, J. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2775-2779 (1975).
43. Lancet, D. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3549-3553 (1976).
44. Zidovetzki, R., Blatt, Y. & Pecht, I. *Biochemistry* **20**, 5011-5018 (1981).
45. Edmundson, A. B., Ely, K. R. & Herron, J. N. *Molec. Immun.* **21**, 561-576 (1984).
46. Ely, K. R. *et al. Biochemistry* **17**, 158-167 (1978).
47. Jackson, D. C. & Webster, R. G. *Virology* **123**, 69-77 (1982).
48. Mandel, B., *Virology* **69**, 500-510 (1976).
49. Lubeck, M. & Gerhard, W. *Virology* **118**, 1-7 (1982).
50. Crumpton, M. J. *Biochem. J.* **100**, 223-232 (1966).
51. Arnon, R. *The Antigens* Vol. 1 (ed. Sela, M.) 87-159 (Academic, New York, 1973).
52. Dimmock, N. J. *J. gen. Virol.* **65**, 1015-1022 (1984).

53. Srinivasappa, J. *et al. J. Virol.* **57**, 397-401 (1986).
54. Tonegawa, S. *Sci. Am.* **253**, 104-113 (1985).
55. Barth, R. K. *et al. Nature* **316**, 517-523 (1985).
56. Arden, B., Klotz, J., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
57. Gorman, J., Hunkapillar, T. & Hood, L. E. *Cell* **45**, 475-484 (1986).
58. Amyard-Henry, M. *et al. Bull. WHO* **48**, 199-202 (1973).
59. Air, G. M. *Virology* **97**, 468-472 (1979).
60. Wang, B. C. *Meth. Enzym.* **115**, 90-112 (1985).
61. Sussman, J. L. in *Methods and Applications in Crystallographic Computing* (eds Hall, S. R. & Ashida, T.) 206-237 (Oxford University Press, 1984).
62. Hendrickson, W. A. & Konnert, J. H. in *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (ed. Srinivasan, R.) 43-57 (Pergamon, Oxford, 1981).

LETTERS TO NATURE

Non-cosmological redshifts of spectral lines

Emil Wolf

Department of Physics and Astronomy, and Institute of Optics,
University of Rochester, Rochester, New York 14627, USA

We showed in a recent report¹ (see also refs 2-4) that the normalized spectrum of light will, in general, change on propagation in free space. We also showed that the normalized spectrum of light emitted by a source of a well-defined class will, however, be the same throughout the far zone if the degree of spectral coherence of the source satisfies a certain scaling law. The usual thermal sources appear to be of this kind. These theoretical predictions were subsequently verified by experiments⁵. Here, we demonstrate that under certain circumstances the modification of the normalized spectrum of the emitted light caused by the correlations between the source fluctuations within the source region can produce redshifts of spectral lines in the emitted light. Our results suggest a possible explanation of various puzzling features of the spectra of some stellar objects, particularly quasars.

To explain why source correlations influence the spectrum of the emitted light consider a very simple example. Suppose that two point sources P_1 and P_2 have identical spectra $S_Q(\omega)$ and that measurements on the emitted field are made at some point P . The sources are assumed at rest relative to an observer at P . Assuming that the source fluctuations can be described by a stationary ensemble, the field at P may be characterized by an ensemble $\{V(P, \omega)\}$ of frequency-dependent realizations⁶, each of the form

$$V(P, \omega) = Q(P_1, \omega) \frac{e^{ikR_1}}{R_1} + Q(P_2, \omega) \frac{e^{ikR_2}}{R_2} \quad (1)$$

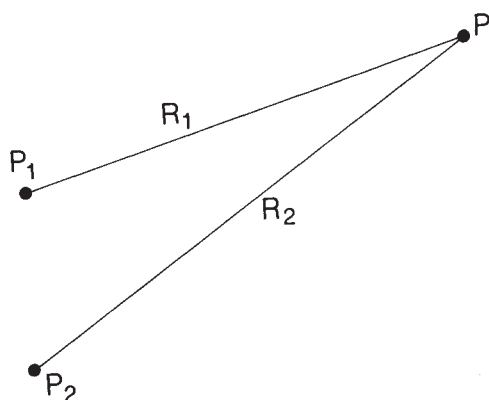


Fig. 1 Illustrating the notation relating to derivation of the formula (4).

where $\{Q(P_j, \omega)\}$, ($j = 1, 2$), characterize the strengths of the two fluctuating point sources, R_1 and R_2 are the distances from P_1 to P and from P_2 to P respectively (see Fig. 1) and $k = \omega/c$, c being the speed of light *in vacuo*. For simplicity polarization effects are ignored and hence V and Q are taken to be scalars. The spectrum of the light at P is then given by

$$S_V(P, \omega) = \langle V^*(P, \omega) V(P, \omega) \rangle \quad (2)$$

where the asterisk denotes the complex conjugate and the angular brackets denote the ensemble average. On substituting from equation (1) into equation (2) and using the fact that

$$\langle Q^*(P_1, \omega) Q(P_1, \omega) \rangle = \langle Q^*(P_2, \omega) Q(P_2, \omega) \rangle = S_Q(\omega) \quad (3)$$

the following expression is obtained for the spectrum of the emitted light at P :

$$S_V(P, \omega) = \left(\frac{1}{R_1^2} + \frac{1}{R_2^2} \right) S_Q(\omega) + \left[W_Q(P_1, P_2, \omega) \frac{e^{ik(R_2 - R_1)}}{R_1 R_2} + \text{c.c.} \right] \quad (4)$$

Here

$$W_Q(P_1, P_2, \omega) = \langle Q^*(P_1, \omega) Q(P_2, \omega) \rangle \quad (5)$$

is the so-called cross-spectral density of the source fluctuations and c.c. denotes the complex conjugate.

The formula (4) shows that the spectrum $S_V(P, \omega)$ is, in general, not just proportional to $S_Q(\omega)$ but is modified by the correlation, characterized by $W_Q(P_1, P_2, \omega)$, between the fluctuations of the two source strengths $Q(P_1, \omega)$ and $Q(P_2, \omega)$. Only in some very special cases, for example, when the source fluctuations are uncorrelated [$W_Q(P_1, P_2, \omega) = 0$] will $S_V(P, \omega)$ be proportional to $S_Q(\omega)$. Hence, in general, the spectrum of the light generated by two point sources depends not only on their spectra but also on the correlation between the fluctuations of their strengths.

A generalization of the elementary formula (4) for radiation from three-dimensional steady-state (that is, statistically stationary) sources of any state of coherence is known⁷. Of special interest in the present context is the form that the formula takes when the source has the same normalized spectrum $s_Q(\omega)$, ($\int_0^\infty s_Q(\omega) d\omega = 1$) at each point in the source region and has a degree of spectral coherence³ (appropriately normalized cross-spectral density) $\mu_Q(\mathbf{r}_1, \mathbf{r}_2, \omega)$ that depends on the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of any source points P_1 and P_2 only through their difference $\mathbf{r}_2 - \mathbf{r}_1$. If, in addition, for each frequency that significantly contributes to the source spectrum, the spectral correlation length [the effective spatial width $|\Delta \mathbf{r}|$ of $|\mu(\mathbf{r}', \omega)|$] is small compared to the linear dimensions of the source, the normalized spectrum $s_V^{(\infty)}(\mathbf{u}, \omega)$ of the emitted light in the far zone, in a direction specified by a unit vector $\mathbf{u}$, becomes (see equation (3.11) of ref. 8)

$$s_V^{(\infty)}(\mathbf{u}, \omega) = \frac{s_Q(\omega) \tilde{\mu}_Q(k\mathbf{u}, \omega)}{\int s_Q(\omega) \tilde{\mu}_Q(k\mathbf{u}, \omega) d\omega} \quad (6)$$

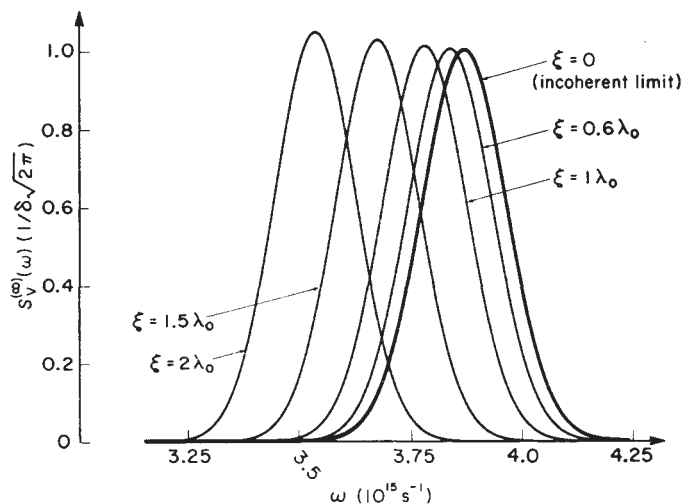


Fig. 2 Spectra $s_V^{(\infty)}(\omega)$ of the far field from sources with spectrum $s_Q(\omega) = (\delta\sqrt{2\pi})^{-1} \exp[-(\omega - \omega_0)^2/2\delta^2]$ and degree of spectral coherence $\mu_Q(\mathbf{r}', \omega) = \exp(-r'^2/2\zeta^2)$, with $\omega_0 = 3.887 \times 10^{15} \text{ s}^{-1}$ ($\lambda_0 = 4,861 \text{ \AA}$) and $\delta = 9.57 \times 10^{13} \text{ s}^{-1}$, for several selected values of the effective source-correlation length ζ . The solid curve ($\zeta \rightarrow 0$) also represents the source spectrum $s_Q(\omega)$.

where $\tilde{\mu}_Q(\mathbf{K}, \omega)$ is the three-dimensional spatial Fourier transform of the degree of spectral coherence $\mu_Q(\mathbf{r}', \omega) \equiv \mu_Q(\mathbf{r}_2 - \mathbf{r}_1, \omega)$.

Let us now choose as the normalized source spectrum $s_Q(\omega)$ a spectral line with a gaussian profile

$$s_Q(\omega) = \frac{1}{\delta\sqrt{2\pi}} \exp[-(\omega - \omega_0)^2/2\delta^2] \quad (\delta \ll \omega_0) \quad (7)$$

and suppose that at each effective frequency ω , the source correlation decreases with the separation $|\mathbf{r}'| = |\mathbf{r}_2 - \mathbf{r}_1|$ of any two source points in a gaussian manner, that is

$$\mu_Q(\mathbf{r}', \omega) = \exp[-r'^2/2\sigma_\mu^2(\omega)] \quad (8)$$

On taking the Fourier transform of equation (8) and substituting the resulting expression into equation (6) we obtain the following expression for the normalized spectrum of the emitted light in the far zone (see equation (3.21) of ref. 8)

$$s_V^{(\infty)}(\omega) = \frac{s_Q(\omega)\sigma_\mu^3(\omega) \exp\{-\frac{1}{2}[k\sigma_\mu(\omega)]^2\}}{\int_0^\infty s_Q(\omega)\sigma_\mu^3(\omega) \exp\{-\frac{1}{2}[k\sigma_\mu(\omega)]^2\} d\omega} \quad (9)$$

Here, $s_V^{(\infty)}(\omega)$ is written in place of $s_V^{(\infty)}(\mathbf{u}, \omega)$, because the spectrum of the far field is now independent of $\mathbf{u}$, as a consequence of the assumed isotropy of μ_Q (see equation (8)).

The formula (9) shows that the spectrum of the emitted light in the far zone depends both on the spectrum of the source fluctuations and on the manner in which the effective source correlation length $\sigma_\mu(\omega)$ depends on the frequency ω .

Let us consider two particular cases. (1) Suppose first that $\sigma_\mu(\omega)$ is independent of ω . Letting ζ denote the (now constant) value of σ_μ and with $s_Q(\omega)$ given by equation (7), one can readily evaluate the integral in the denominator on the right of equation (9) and one then finds that

$$s_V^{(\infty)}(\omega) \approx \frac{\alpha}{\delta\sqrt{2\pi}} \exp\left[-\left(\omega - \frac{\omega_0}{\alpha^2}\right)^2 / 2(\delta/\alpha^2)^2\right] \quad (10)$$

where

$$\alpha^2 = 1 + \left(\frac{\delta}{\zeta}\right)^2 \quad (11a)$$

and

$$\frac{1}{\Delta} = \frac{\zeta}{c} \quad (11b)$$

When the source is effectively spatially incoherent, $\zeta \rightarrow 0$. Then according to equation (11) $\Delta \rightarrow \infty$ and $\alpha \rightarrow 1$ and it follows from equations (10) and (7) that in this case

$$s_V^{(\infty)}(\omega) \rightarrow s_Q(\omega) \quad (12)$$

Hence, in the limiting case of a completely incoherent source of the class that is considered here, the normalized spectrum of

the emitted light in the far zone is identical with the normalized spectrum of the source fluctuations.

However, when the source fluctuations are correlated over an effective distance $\zeta > 0$, equation (10) shows that the spectrum $s_V^{(\infty)}(\omega)$, although it is also a line with a gaussian profile, is centred at a lower frequency $\omega'_0 \approx \omega_0/\alpha^2 < \omega_0$. Hence the source correlations give rise to a spectral line $s_V^{(\infty)}(\omega)$ that is redshifted with respect to the spectral line produced by the completely spatially incoherent source with the source spectrum $s_Q(\omega)$. The shifted line is narrower, having root-mean-square width $\delta' = \delta/\alpha < \delta$ and has α -times greater height. Examples of spectra of light in the far zone, produced by several sources which emit the same spectral line but which have different correlation lengths are shown in Fig. 2. From the formula (10) one can readily deduce that the relative shift of the line, namely,

$$z \equiv \frac{\lambda_0 - \lambda'_0}{\lambda_0} = -\frac{\omega_0 - \omega'_0}{\omega'_0} \quad (13)$$

($\lambda_0 = 2\pi c/\omega_0$, $\lambda'_0 = 2\pi c/\omega'_0$) is given by

$$z = \left(\frac{\delta}{\Delta}\right)^2 = \left(\frac{\delta}{c}\right)^2 \zeta^2 \quad (14)$$

which shows that in this case the redshift increases quadratically with the spectral source-correlation length ζ . (2) Next consider the situation when $\sigma_\mu(\omega) = a/\omega$ where a is a positive constant. The expression (9) for the normalized spectrum of the emitted light in the far zone now reduces to

$$s_V^{(\infty)}(\omega) = \frac{s_Q(\omega)/\omega^3}{\int_0^\infty [s_Q(\omega)/\omega^3] d\omega} \quad (15)$$

Note that this expression is independent of the value of the constant a .

When $s_Q(\omega)$ is a line with a gaussian profile, given by equation (7), the spectrum $s_V^{(\infty)}(\omega)$, given by equation (15) is no longer strictly gaussian but it can be closely approximated by a gaussian and can be shown to be redshifted with respect to $s_Q(\omega)$ by the relative amount

$$z \approx 3\left(\frac{\delta}{\omega_0}\right)^2. \quad (16)$$

An example of this situation is illustrated in Fig. 3.

This case [$\sigma_\mu(\omega) = a/\omega$] is of special interest because, according to equation (8), the degree of spectral coherence is now given by

$$\mu_Q(\mathbf{r}', \omega) = \exp[-(kr')^2/2(a/c)^2], \quad (17)$$

that is, it has the functional form

$$\mu_Q(\mathbf{r}', \omega) = f(kr') \quad (k = \omega/c = 2\pi/\lambda) \quad (18)$$

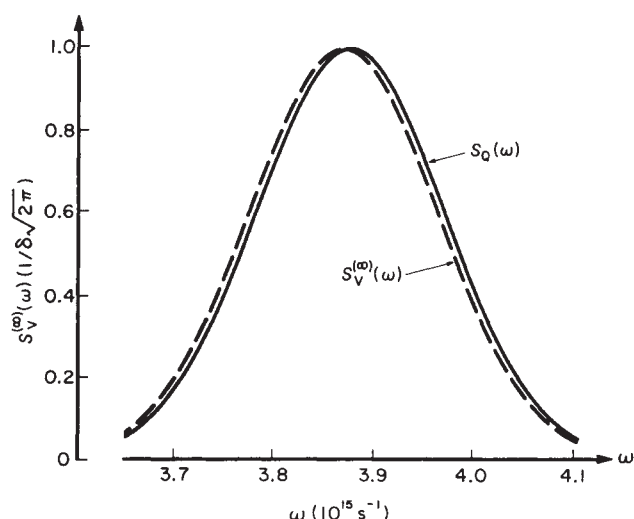


Fig. 3 The spectrum $S_V^{(\infty)}(\omega)$ of the far field from a source with source spectrum $S_Q(\omega) = (\delta\sqrt{2\pi})^{-1} \exp[-(\omega - \omega_0)^2/2\delta^2]$ and degree of spectral coherence $\mu_Q(\mathbf{r}', \omega) = \exp[-(kr')^2/2(a/c)^2]$ (a = an arbitrary constant), with $\omega_0 = 3.887 \times 10^{15} \text{ s}^{-1}$ ($\lambda_0 = 4,861 \text{ \AA}$) and $\delta = 9.57 \times 10^{13} \text{ s}^{-1}$. The source spectrum $S_Q(\omega)$ is shown for comparison. Note that $\mu_Q(\mathbf{r}', \omega)$ now obeys the scaling law.

Thus the degree of spectral coherence of the source distribution now satisfies the three-dimensional analogue of a requirement (called the scaling law) derived in ref. 1, as a sufficient condition for the spectrum of the light emitted by a planar secondary source of a well-defined class to have certain invariance properties on propagation. It will be shown in another publication (J. T. Foley and E. Wolf, in preparation) that for three-dimensional primary sources of an analogous class, whose degree of coherence satisfies this law, the spectrum of the emitted light has similar invariance properties. We conjecture that the usual thermal sources obey such a scaling law.

Now briefly consider the question of a physical mechanism for producing source correlations. Such correlations must clearly be manifestations of some cooperative phenomena. At the atomic level possible candidates may perhaps be superradiance and superfluorescence⁹. An effect of this kind was first predicted by Dicke in 1954 when he showed¹⁰ that under certain circumstances energy from excited atoms may be released cooperatively in a much shorter time than the natural lifetime of the excited states of the atoms and with much larger emission intensity than would be obtained were the atoms radiating independently. However not enough is known at present about the coherence properties of large three-dimensional systems of this kind to make it possible to determine whether superradiance and superfluorescence might involve correlations that could give rise to spectral line shifts.

There is, however, quite a different mechanism, which can be described at the macroscopic level, and which can imitate effects of source correlations; namely effects of correlations between the refractive index at pairs of points in a spatially random but statistically homogeneous, time-invariant medium. If a wave illuminates such a medium, say a dilute gas, then, as is well known, the medium acts as a secondary source, namely as a set of oscillating charges set in motion by the incident wave. The secondary waves produced by the oscillating charges then combine with each other and with the incident wave and generate the scattered field. If the gas is not too dilute the collective response of the microscopic charges to the incident field can be described by macroscopic parameters such as the dielectric susceptibility or the refractive index. Now within the accuracy of the first Born approximation the basic equation for scattering

is of the same form as the basic equation for radiation from primary sources, the 'equivalent source' for scattering being the product of the scattering potential (which is a simple function of the refractive index) and of the incident wave. This correspondence clearly implies that our results regarding the effects of source correlations on the spectrum of the emitted light must have analogues regarding the effects of a spatially random medium with correlated refractive index distribution on the spectrum of the light that is scattered by it. This topic will be discussed elsewhere.

Let us now consider some implications of this analysis. Using equation (14), the spectral line in Fig. 2, produced by the source whose correlation length $\zeta = \lambda_0$ is readily found to have a redshift given by $z = 0.0241$ with respect to the source spectrum. It is of interest to note that if an observer detected such a redshift unaware of its true origin and interpreted it on the basis of the Doppler shift formula $v/c = \Delta\lambda/\lambda_0 = z$ he would incorrectly conclude that the source was receding from him with a speed $v = 0.0241c \approx 7,230 \text{ km s}^{-1}$.

It seems worthwhile to note that there is a maximum line shift that can be produced by source correlations. This can be seen from the basic formula (6) which indicates that $S_V^{(\infty)}(\mathbf{u}, \omega) = 0$ when $S_Q(\omega) = 0$, implying that the spectrum of the far field can only contain those frequencies that are already present in the source spectrum. Consequently the maximum attainable frequency shift of the line cannot exceed its effective frequency range. However, any frequency contribution from the source spectrum to the normalized spectrum of the far field can be greatly magnified or greatly reduced, as is evident from equation (6) and from Fig. 2.

We have mainly considered effects of source correlations under circumstances when the source spectrum consists of a single line and when the degree of spectral coherence μ_Q that characterizes the source correlations depends on a single parameter. Preliminary calculations show that with a suitably chosen μ_Q which depends on a larger number of parameters, redshifts of several lines may be produced, all of which will have approximately the same z -values.

In this article we have considered redshifts of spectral lines. However, it is not difficult to specify source correlations which will produce blueshifts. Examples of this kind are given in a forthcoming publication¹¹.

It seems plausible that the mechanism discussed in this article may be responsible for some of the so far unexplained features of quasar spectra, including line asymmetries and small differences in the observed redshifts of different lines. In this connection it is of interest to recall that the role of coherence in the emission of radiation from quasars was stressed by Hoyle, Burbidge and Sargent in a well-known article¹².

I thank Mr A. Gamliel and Mr K. Kim for carrying out computations relating to the analysis presented in this article. The fact that scattering can also produce shifts of spectral lines was noted independently by Professor Franco Gori, who informed me of this result when commenting on an early version of the manuscript of this article. This investigation was supported by the NSF and by the US Air Force Geophysical Laboratory.

Received 10 September 1986; accepted 3 February 1987.

1. Wolf, E. *Phys. Rev. Lett.* **56**, 1370-1372 (1986).
2. Mandel, L. *J. opt. Soc. Am.* **51**, 1342-1350 (1961).
3. Mandel, L. & Wolf, E. *J. opt. Soc. Am.* **66**, 529-535 (1976).
4. Gori, F. & Grella, R. *Optics Commun.* **49**, 173-177 (1984).
5. Morris, G. M. & Faklis, D. *Optics Commun.* (in the press).
6. Wolf, E. *J. opt. Soc. Am.* **72**, 343-351 (1982); *J. opt. Soc. Am. A* **3**, 76-85 (1986).
7. Carter, W. H. & Wolf, E. *Optica Acta* **28**, 227-244 (1981).
8. Carter, W. H. & Wolf, E. *Optica Acta* **28**, 245-259 (1981).
9. Schuurmans, M. F. H., Vrehen, Q. H. F., Polder, D. & Gibbs, H. M. in *Advances in Atomic and Molecular Physics* Vol. 17 (eds Bates, D. & Bederson, B.) 167-228 (Academic, New York, 1981).
10. Dicke, R. H. *Phys. Rev.* **93**, 99-110 (1954).
11. Wolf, E. *Optics Commun.* (in the press).
12. Hoyle, F., Burbidge, G. R. & Sargent, W. L. W. *Nature* **209**, 751-753 (1966).

Speculations about the origin of H_3O^+ seen in cometary mass spectra

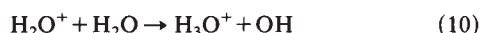
Edmond Murad* & Peter Bochsler†

* Air Force Geophysics Laboratory, Hanscom AFB, Maryland 01731, USA

† Physikalisches Institut, Universität Bern, Sidlerstrasse 5, CN-3012 Bern, Switzerland

Recent mass spectrometer measurements of the composition of the coma of Giacobini-Zinner and Halley comets have indicated the presence of large amounts of H_3O^+ . Here, we suggest that some of these ions arise from the photoionization of the water dimer, which is formed by direct vaporization of ice from solid surfaces at the cometary nucleus or from ice particles ejected in cometary jets.

The recent space missions ICE (International Cometary Explorer) to comet Giacobini-Zinner, and Giotto, Sakigake, Suisei and Vega to comet Halley have yielded data concerning the ion and neutral compositions of the gaseous clouds surrounding these comets. The ICE data indicate the presence of ions in the mass range $16 \leq m/q \leq 19$ (where m is the mass and q is the charge) which are interpreted to be due to O^+ , OH^+ , H_2O^+ and H_3O^+ , respectively^{1,2}. The Giotto mission has yielded similar, but more detailed, results^{3,4}. The origin of these ions is generally understood in terms of known chemical processes and reactions^{5,6}, for example, H_2O^+ is produced from the photoionization of H_2O , and O^+ and OH^+ undoubtedly arise from H_2O undergoing simple photolytic reactions followed by photoionization⁵. The observation of H_3O^+ by the ICE and Giotto missions has been taken as evidence for the process (reaction 10 in Table 1)



occurring in the coma, where the concentration of H_2O is highest. At thermal energies (0.025 eV), this reaction is very fast⁷, its rate

Table 1 Kinetic scheme used for dimer evaluation

Reaction	Rate coefficient k
1. $\text{H}_2\text{O} + h\nu \rightarrow \text{H} + \text{OH}$	$1 \times 10^{-5*}$
2. $\text{H}_2\text{O} + h\nu \rightarrow \text{H}_2 + \text{O}$	$1.3 \times 10^{-6*}$
3. $\text{H}_2\text{O} + h\nu \rightarrow \text{H}^+ + \text{OH} + e$	$1.3 \times 10^{-8*}$
4. $\text{H}_2\text{O} + h\nu \rightarrow \text{OH}^+ + \text{H} + e$	$5.5 \times 10^{-8*}$
5. $\text{H}_2\text{O} + h\nu \rightarrow \text{O}^+ + \text{H}_2 + e$	$5.9 \times 10^{-9*}$
6. $\text{H}_2\text{O} + h\nu \rightarrow \text{H}_2\text{O}^+ + e$	$3.3 \times 10^{-7*}$
7. $(\text{H}_2\text{O})_2 + h\nu \rightarrow 2\text{H}_2\text{O}$	$1 \times 10^{-4§}$
8. $(\text{H}_2\text{O})_2 + h\nu \rightarrow (\text{H}_2\text{O})_2^+ + e$	$1.6 \times 10^{-7†}$
9. $(\text{H}_2\text{O})_2 + h\nu \rightarrow \text{H}_3\text{O}^+ + \text{OH} + e$	$6 \times 10^{-7†}$
10. $\text{H}_2\text{O}^+ + \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}^+ + \text{OH}$	$2 \times 10^{-9*}$
11. $(\text{H}_2\text{O})_2^+ + \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}_2^+ + \text{OH}$	$1 \times 10^{-9‡}$
12. $\text{H}_2\text{O}^+ + \text{H}_2\text{O} \rightarrow \text{H}_2\text{O}^+ + \text{H}_2\text{O} + e$	$1 \times 10^{-10§}$
13. $\text{H}^+ + (\text{H}_2\text{O})_2 \rightarrow \text{H}_3\text{O}^+ + \text{H}_2\text{O}$	$5 \times 10^{-10§}$
14. $\text{H}^+ + (\text{H}_2\text{O})_2 \rightarrow [(\text{H}_2\text{O})_2^+ + \text{H}] \rightarrow \text{H}_3\text{O}^+ + \text{OH} + \text{H}$	$1 \times 10^{-9§}$
15. $\text{H}^+ + (\text{H}_2\text{O})_2 \rightarrow \text{H}_3\text{O}^+ + \text{H}_2\text{O}$	$3 \times 10^{-9§}$
16. $\text{H}_3\text{O}^+ + h\nu \rightarrow \text{H}^+ + \text{H}_2\text{O}$	$5 \times 10^{-6§}$
17. $\text{H}_3\text{O}_2^+ + h\nu \rightarrow \text{H}_3\text{O}^+ + \text{H}_2\text{O}$	$1 \times 10^{-5§}$
18. $\text{H}_3\text{O}^+ + \text{H}_2\text{O} \rightarrow \text{H}^+ + \text{H}_2\text{O} + \text{H}_2\text{O}$	$5 \times 10^{-11§}$
19. $\text{H}_3\text{O}_2^+ + \text{H}_2\text{O} \rightarrow \text{H}_3\text{O}^+ + \text{H}_2\text{O} + \text{H}_2\text{O}$	$1 \times 10^{-10§}$
20. $\text{H}^+ + \text{H}_2\text{O} \rightarrow \text{H}_2\text{O}^+ + \text{H}$	$8 \times 10^{-9¶}$

k_1 – k_9 , k_{16} and k_{17} are in units of s^{-1} . Others are in units of $\text{cm}^3 \text{molec}^{-1} \text{s}^{-1}$.

* Huebner and Giguere (ref. 5).

† Total photoionization rate of dimer is assumed to be twice that of the monomer (see text).

‡ Good *et al.* (ref. 13).

§ Estimated.

¶ Huntress *et al.* (ref. 19).

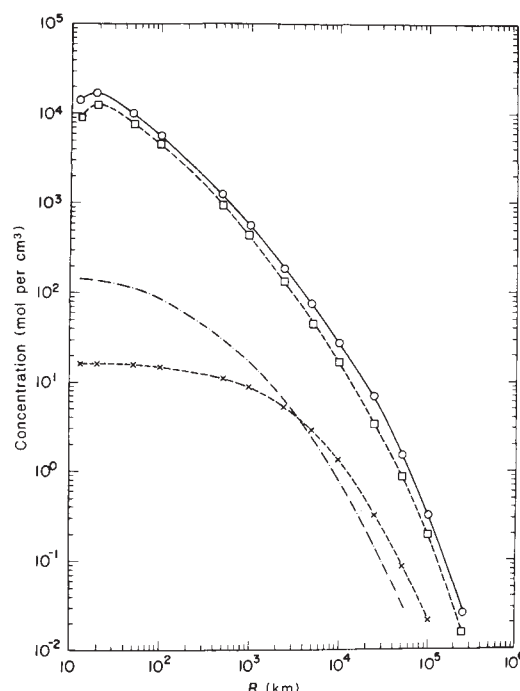


Fig. 1 The effect of dimer on the concentration of H_3O^+ using the reaction scheme shown in Table 1. R , distance from centre of comet. $\circ$, H_3O^+ and 10% dimer; $\square$, H_3O^+ and no dimer; $\times$, $(\text{H}_2\text{O})_2^+$ and 10% dimer; $\cdots$, H_3O_2^+ and 10% dimer.

constant, k_{10} being $1.7 \times 10^{-9} \text{ cm}^3 \text{ molec}^{-1} \text{ s}^{-1}$. Because this reaction is exothermic, however, the cross section, σ , decreases with increasing collision energy. σ_{17} decreases from 60×10^{-16} at 0.2 eV to $\sim 1.5 \times 10^{-16} \text{ cm}^2$ at 100 eV (ref. 8).

It is generally assumed that comets consist of ice which is evaporated as the comet nears the Sun. Thus the gaseous species are formed by the vaporization of ice. Studies of the vaporization of ice indicate that some water dimer, $(\text{H}_2\text{O})_2$, is always formed. In this report we consider the evaporation of water from the comet nucleus and dust and evaluate the effect of dimer on the expected ionic species.

A number of investigators have studied the thermodynamics of vaporization of water^{9,10}. These studies indicate that the equilibrium concentration of the dimer can be substantial, particularly at low temperatures. For example, using the data provided by Slanina¹⁰, we calculate that at 200, 250, 300 and 400 K the mole fraction of the dimer is 0.3, 0.12, 0.05 and 0.01, respectively. The data of Balsiger *et al.*³ indicate that the temperature at the nuclear surface is about 300 K. Thus the dimer concentration would be 5% if the vaporization were to proceed in an equilibrium manner. However, this is not the case. Pictures obtained by Giotto¹¹ indicate the presence of two beams of dust and molecules effusing from the nucleus of the comet, indicating, perhaps, a supersonic expansion through holes. If that is the case, then isentropic expansion of the gas through the hole would lead to cooling and further condensation of the water molecules. We assume for the purpose of this exercise that the evaporation temperature is 250 K, that is, that the dimer fraction is at least 0.12.

Following vaporization, the dimers are photoionized. Laboratory studies of the photoionization and fragmentation of the water dimer¹² indicate that the primary process yields the parent ion (reaction 8 in Table 1) at $11.21 \pm 0.09 \text{ eV}$. Fragmentation of the parent ion to H_3O^+ (reaction 9 in Table 1) is observed at $11.73 \pm 0.03 \text{ eV}$. The photoionization studies mentioned above and mechanistic studies of ion-neutral collisions¹³ have indicated that the dimer ion, $(\text{H}_2\text{O})_2^+$, may, in fact, be thought of as $\text{H}_3\text{O}^+ \cdot \text{OH}$. In any case, at $\lambda = 100 \text{ nm}$ (12.4 eV), reaction 9 is 3–4 times as large as reaction 8.¹² Thus at first glance it is

reasonable to suppose that some of the observed H_3O^+ arises from the dimer through reaction 9. It is worth noting that the thresholds for reactions 8 and 9 are both lower than the ionization potential of water (12.61 eV). The studies which have been published about the photoionization of the dimer do not provide absolute cross-sections. It seems reasonable to assume that the photoionization cross-section of the dimer is twice that of the monomer, at least to a first approximation. In addition, because the dimer has a lower ionization potential than the monomer, there will be an enhancement in the photoionization because of increased solar flux at longer wavelengths. Huebner and Giguere⁵ quote an ionization rate for H_2O of $3.34 \times 10^{-7} \text{ s}^{-1}$. From this we estimate the ionization rate for the dimer to be $1 \times 10^{-6} \text{ s}^{-1}$.

Within the magnetopause, ions and neutrals collide at thermal energies³. Outside the magnetopause, the interaction is much more complex. The densities are quite low, and collisions, when they occur, do so at energies varying between 0.2 eV and 35 keV; at high energies the cross-section for reaction 10 will be quite small. The expansion velocity of neutral H_2O has been determined for the case of Halley by Krankowsky *et al.*⁴ to be $900 \pm 200 \text{ m s}^{-1}$.

Using the reaction network described in Table 1 and the very simple model of a radially outgassing body at a moderate temperature in the unattenuated photon flux from the Sun at 1 AU we have calculated radial profiles of concentrations of H_2O^+ , H_3O^+ , $(\text{H}_2\text{O})_2^+$ and $(\text{H}_5\text{O}_2)^+$ (a production rate for H_2O of 5×10^{29} molecules per second and a constant radial expansion velocity of 1 km s^{-1} were assumed). The intensities of these ions are shown in Fig. 1. The dashed line indicates the H_3O^+ concentration if no dimer is assumed, and the full line indicates the intensity for the outgassing of 10% dimer ($Q_{\text{dimer}} = 5 \times 10^{-28}$ molecules per second). The assumption of unattenuated photon flux is not valid for the innermost regions of the coma whereas the reactions outside the cometopause will occur at much higher energies than assumed for our model. Hence it is clear that this model can only give an impressionistic picture of composition in the intermediate range.

It appears that on the average, the presence of 10% dimer enhances the concentration of H_3O^+ by 25–30%; simultaneously, the concentration of H_2O^+ is increased 7–20% depending on the distance from the nucleus. This result of our simple model calculation can be understood as a consequence of the almost double ionization yield of reactions 8 and 9 compared to reactions 3–6 which produce ions from the monomer. Reaction 12, in turn, adds to the ionized fraction in a nonlinear manner. Our result is consistent with suggestions by Coplan *et al.*¹⁴ on the possible importance of carbon clusters in the coma of Giacobini–Zinner and observations of Eberhardt *et al.*¹⁵ and the discussions of Wallis *et al.*¹⁶ which clearly demonstrate that volatile compounds, clusters and organic sub-micron grains could severely influence the chemical balance in cometary comae.

Finally, we wish to point out that we have attempted in this work to consider sound thermochemical data and to include them into a consideration of the chemical processes which occur in cometary atmospheres. This is a preliminary treatment of the subject and more quantitative treatment will have to be performed later. A comprehensive study should include all the reactions considered in a recent review¹⁷ as well as reactions of electrons with electronegative molecules such as H_2O and CO_2 . The latter molecules yield OH^- , H^- and O^- upon collision with low energy electrons¹⁸.

P.B. wishes to thank the Swiss National Science Foundation for financial support; he also wishes to thank USAF for a Window on Science grant.

Received 17 November 1986; accepted 15 January 1987.

- Ogilvie, K. W., Coplan, M. A., Bochsler, P. & Geiss, J. *Science* **232**, 374–377 (1986).
- Coplan, M. A., Ogilvie, K. W., A'Hearn, M. F., Bochsler, P. & Geiss, J. *J. geophys. Res.* **92**, 39–46 (1987).

- Balsiger, H. *et al. Nature* **321**, 330–334 (1986).
- Krankowsky, D. *et al. Nature* **321**, 326–329 (1986).
- Huebner, W. F. & Giguere, P. T. *Astrophys. J.* **238**, 753–762 (1980).
- Huebner, W. F. in *The Photochemistry of Atmospheres* 437–481 (Academic, New York 1985).
- Albritton, D. L. *At. Data nucl. Tabl.* **22**, 1–101 (1978).
- Lishawa, C. R., Salter, R. H. & Murad, E. *J. chem. Phys.* (in the press).
- Curtiss, L. A., Frurip, D. J. & Blander, M. *J. chem. Phys.* **71**, 2703–2711 (1979).
- Slanina, Z. *Chem. Phys. Lett.* **127**, 67–72 (1986).
- Keller, H. U. *et al. Nature* **321**, 320–326 (1986).
- Ng, C. Y. *et al. J. chem. Phys.* **67**, 4235–4237 (1977).
- Good, A., Durden, D. A. & Kebarle, P. *J. chem. Phys.* **52**, 222–229 (1970).
- Coplan, M. A., Ogilvie, K. W. & A'Hearn, M. F. *Pap. presented at 20th ESLAB Symposium Exploration of Halley's Comet*, Heidelberg, 27–31 October 1986.
- Eberhardt, P. *et al. Pap. presented at 20th ESLAB Symposium Exploration of Halley's Comet*, Heidelberg, 27–31 October 1986.
- Wallis, M. K., Rabilizirov & Wickramasinghe, N. C. *Pap. presented at 20th ESLAB Symposium Exploration of Halley's Comet*, Heidelberg, 27–31 October 1986.
- Gombosi, T. I., Nagy, A. F. & Cravens, T. E. *Rev. Geophys.* **24**, 667–700 (1986).
- Christophorou, L. G., McCorkle, D. L. & Christodoulides, A. A. in *Electron Molecule Interactions and Their Applications* Vol. 1 (ed. Christophorou, L. G.) 477–619 (Academic, New York, 1984).
- Huntress, W. T., Kim, J. K. & Theard, L. P. *Chem. Phys. Lett.* **29**, 189–190 (1974).

Monopoles in the rotating superfluid helium-3 A–B interface

M. M. Salomaa

Low Temperature Laboratory, Helsinki University of Technology, SF-02150 Espoo, Finland

The ^3He A–B interface provides an unprecedented type of superfluid boundary between two degenerate macroscopically coherent quantum systems which display different broken symmetries. As a consequence, the rotating superfluid A–B phase boundary is here predicted to exhibit novel physical features: vortices in superfluid ^3He cannot perforate as such across the A–B interface. Instead, at the rotating phase boundary, the macroscopic rotation condition is fulfilled by a lattice of monopoles—or by exotic vortices with half-integer circulation quanta. Here we suggest that the resulting areal density of monopoles which persist in the phase boundary possibly enables, for the first time, their experimental detection in superfluid ^3He .

The superfluid phase boundary between ^3He -A and ^3He -B also serves to provide a unique analogy to the cosmic domain walls^{1,2} that are believed to have precipitated between regions of the early Universe³ in different vacuum states. Hence the proposed experiments on the superfluid ^3He A–B interface would constitute an important analogy model to phenomena at cosmic domain walls, where monopoles are believed to occur on the postulated cosmic strings⁴ at the event horizon of the Universe. Another intriguing 'cosmological' experiment was recently suggested⁵ on superfluid ^4He ; the latter proposal is to cool liquid ^4He suddenly through the superfluid transition whereby vorticity may evolve at the interfaces of the superfluid condensation fronts. An analogy model experiment on superfluid ^3He would be potentially more rewarding⁶, because for example its vortex density could be probed directly with use of the NMR (nuclear magnetic resonance) technique. Instead of such an essentially chaotic scenario, we support here controlled measurements on the A–B interface, also with inherent cosmological significance.

Recent experimental and theoretical investigations of rotating superfluid ^3He have served to identify several unusual types of new linear order-parameter defects—vortices^{6,7}—both in superfluid ^3He -A and in ^3He -B. The purpose of this report is to propose such unique and quite specific experimental circumstances under which pointlike order-parameter defect structures, monopoles, will exist in sufficient quantity to be detected in practical experiments; rotation proves indispensable for the method.

Monopoles were first suggested by Dirac⁸ in a mathematical description of isolated magnetic charges. These pointlike objects would be singularities of the electromagnetic field. Such

magnetic charges possess dual character. On one hand, the Dirac monopole is a point singularity in the distribution of magnetic field $\mathbf{H}$ directed radially from the centre of the charge; on the other hand, this point defect is connected to a linear singularity ('filament') of the magnetic vector-potential field, $\mathbf{A}$. The filament (in the present context, a quantized vortex line) extends from the centre of the monopole to infinity.

Since the experimental search for magnetic monopoles (charge singularities in field theory) has so far yielded no definite results⁹, there is an incentive to study the formation of pointlike topological structural analogues of monopoles (order-parameter singularities) in different condensed-matter systems. In particular, monopoles were theoretically suggested to be energetically stable in liquid crystals¹⁰; such structures are observed experimentally¹¹.

Monopole-like order-parameter textures in bulk superfluid ^3He have been considered by several authors¹²⁻¹⁹. (Here the analogue of $\mathbf{H}$ is the vorticity of superflow, $\nabla \times \mathbf{v}_s$.) Among the suggested experimental situations under which monopoles have been proposed to occur, we quote: (1) monopole-antimonopole pairs stabilized by negative ions^{12,14,15,18}, (2) dynamic processes of vortex transitions^{7,17} (for example, in $^3\text{He-A}$, continuous doubly quantized vortex $\rightarrow$ two singly quantized singular vortices), and (3) monopole structures on the surface of a sphere^{14,15}. However, to date there exists no evidence for the occurrence of monopoles in ^3He . Clearly, in the proposed schemes, either the experimental circumstances are prohibitively hard to realize, or the density of monopoles obtained in them is not sufficient. In particular, it is difficult to detect moving monopoles during transient conditions.

The superfluid interface between the A and B phases can be stabilized and investigated in detail^{20,21}. The velocity of the moving A-B phase boundary has been measured²² and also calculated theoretically²³. The stationary A-B interface has been investigated previously²⁴. Recently, we have shown²⁵ that the moving A-B interface possesses peculiar and interesting measurable physical properties, for example, a spontaneous anisotropic magnetization, spontaneously flowing supercurrents, such that it may be thought of as a 'vortex sheet', and a phase transition.

It is interesting to examine the rotating A-B interface. For simplicity, consider the angular velocity of rotation Ω perpendicular to the plane of the phase boundary. Note that vortices in $^3\text{He-B}$ possess a non-rotating A-phase-like superfluid core²⁶. On approaching the B $\rightarrow$ A phase boundary, the superfluid core of the $^3\text{He-B}$ vortex grows abruptly, resulting in interfacial 'wetting'. While the A-phase component inside the $^3\text{He-B}$ vortex core is non-rotating, the latter must be spun into rotation in the bulk A phase; this results in an ambiguity at the A-B interface. (In contrast, vortices can pierce the superfluid A-A₁ phase boundary since the order-parameter symmetry group for $^3\text{He-A}_1$ is a subgroup of that for $^3\text{He-A}$.)

The vortex lattice in bulk $^3\text{He-B}$ fulfils the solid-body-like rotation condition, and the vortex lattice in $^3\text{He-A}$ correspondingly obeys it on the opposite side of the B $\rightarrow$ A phase boundary, in the bulk A phase. However, in the plane of the A-B interface, the rotation condition is fulfilled through the formation of a lattice of monopoles, which serves to resolve the dilemma alluded to above. These monopoles can occur as free ends of the vortex lines in $^3\text{He-A}$ or $^3\text{He-B}$, which need not perforate the phase boundary, see Fig. 1a, or as discontinuities on the vortex lines, see Fig. 1b. The monopoles could be studied with help of transverse continuous-wave NMR; their characteristic signature would be observed in the absorption spectrum. (The NMR equation⁷ resembles the Dirac equation⁸.)

Such monopoles in the A-B phase boundary correspond to the 'beads' postulated to be attached to strings²⁷ at a boundary connecting two regions with degenerate ground states. We emphasize that here—unlike for other situations where monopoles in superfluid ^3He have been suggested—there results a dense and stationary array of monopoles on the phase bound-

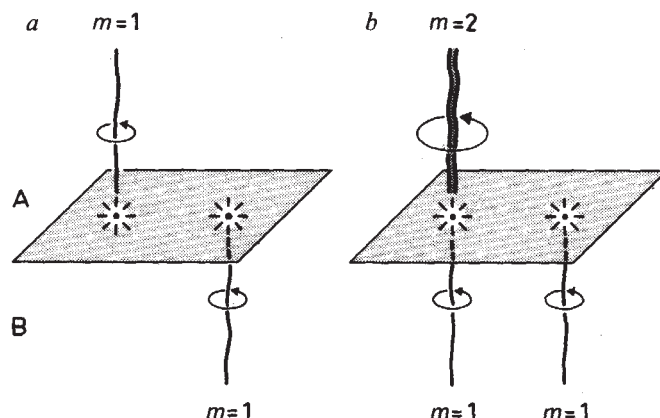


Fig. 1 The vortex array in superfluid ^3He cannot propagate unmodified across the A-B phase boundary. At the interface (hatched), the macroscopic rotation condition is met by a lattice of monopoles; such monopoles provide seeds where singular vortices may preferentially nucleate in $^3\text{He-A}$. These models involve vortices with free ends at the A-B interface; m denotes the number of superflow circulation quanta. *a*, Vortices in $^3\text{He-A}$ and $^3\text{He-B}$ may be disjoint. *b*, A doubly-quantized continuous vortex in $^3\text{He-A}$ cannot propagate through the phase boundary: at least unit quantum of circulation will extinguish at a monopole, while the other quantum may survive to perforate the phase boundary into a $^3\text{He-B}$ vortex. To conserve circulation, another singly quantized $^3\text{He-B}$ vortex line must have a free end on the B $\rightarrow$ A interface.

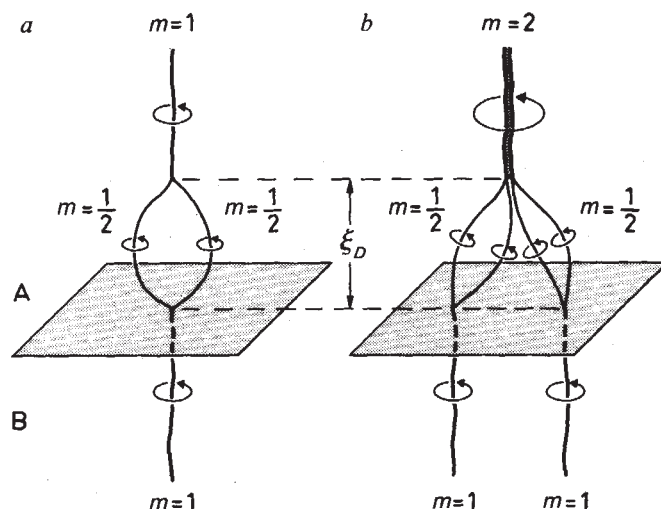


Fig. 2 Vortices with half-integer circulation quantum may be uncovered in the dipole-unlocked 'soft' core (with width on the order of the dipole length, $\xi_D \sim 6 \mu\text{m}$) of the rotating A $\rightarrow$ B phase boundary. *a*, A pair of bare half-quantum vortices couples singly quantized vortices in $^3\text{He-A}$ and $^3\text{He-B}$. *b*, A continuous doubly quantized vortex in $^3\text{He-A}$ connects a pair of vortices in $^3\text{He-B}$ through a quartet of half-quantum vortices.

dary, thus possibly enabling their observation. It would be especially informative to carry out NMR imaging experiments (NMR 'tomography') of rotating superfluid ^3He and the A-B phase boundary in rotation to 'see' the vortex lattices—because of the magnetic anisotropy in the vortex-core regions—and the free ends of vortices, thus proving the existence of monopoles on vortex lines in the A-B interface.

The bulk superfluid A phase is at low magnetic fields in the so-called dipole-locked situation ($d \parallel l$, for a discussion of these spin and orbital anisotropy axes see, for example, ref. 20). Under this condition, the occurrence of half-quantum vortices is prohibited²⁸. However, the A-B phase boundary spontaneously

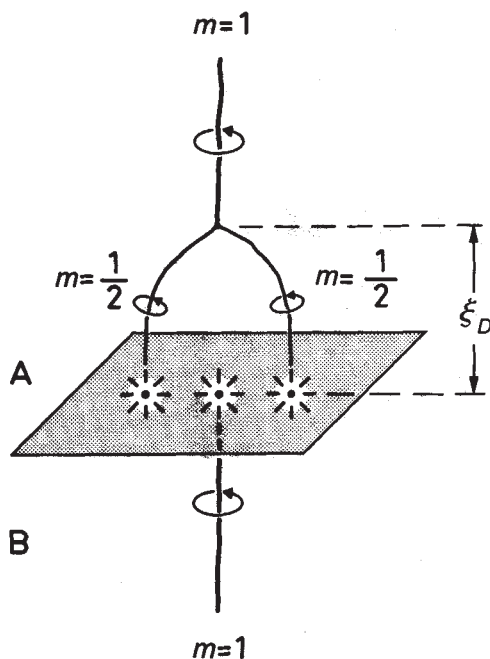


Fig. 3 A possible scheme exhibitive of both half-quantum vortices and monopoles at the A-B interface: The $^3\text{He-A}$ vortex core spontaneously bifurcates into a pair of half-quantum vortices terminating in the A-B boundary.

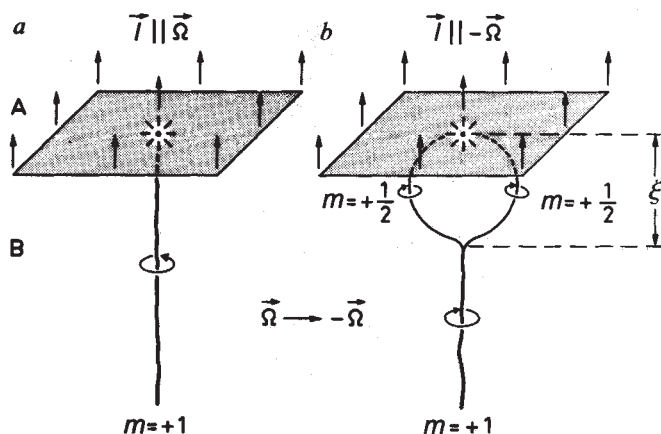


Fig. 4 A peculiar Ω -dependent effect results at the rotating A-B boundary when the orbital anisotropy axis l in $^3\text{He-A}$ is orientated normal to the plane of the phase boundary. *a*, For $\Omega \parallel l$, an axisymmetric monopole on the axisymmetric v vortex in $^3\text{He-B}$ is allowed by symmetry. *b*, With $\Omega \parallel -l$, the free end of the vortex line becomes nonaxisymmetric by necessity. Evidently, there awaits a rich abundance of different types of monopoles to be discovered at the rotating A-B interface; moreover, their quantum state may quite simply be switched, for example, upon reversal of Ω .

reorients the A phase²⁵, such that dipole unlocking occurs ($d \perp l$) in its vicinity. This leads to the formation of a 'soft' core for the A $\rightarrow$ B interface, which is adjustable in size and orientation with an applied magnetic field. The half-quantum vortices may appear in this region. Hence this leads to other exciting scenarios that are also possible at the rotating $^3\text{He A-B}$ interface, involving topologically confined pairs of vortices in $^3\text{He-A}$ with half-integer circulation quanta, see Figs 2 and 3. The occurrence of half-quantum vortices could again be resolved with the help of NMR; their NMR signature is peculiar due to an analogy²⁹ with the Aharonov-Bohm effect.

If 'cosmological' experiments⁵ were performed in superfluid ^3He , they would be most interesting on the A-B coexistence curve, owing to the two different possible degenerate ground states for the system. Unlike for ^4He , due to the supercurrents generated at the A-B phase boundaries²⁵, there prevail additional reasons for spontaneous vorticity to emerge. Nevertheless, we emphasize the power of controlled measurement schemes, and introduce a new method to search for monopoles in superfluid ^3He . Instead of just one pair of moving pointlike objects (as in dynamic vortex phenomena), the present scheme involves monopoles in profusion at the rotating phase boundary—with superficial density not less than the areal number density of vortices—thus possibly enabling their experimental verification.

Quantized vorticity in the rotating superfluid A-B phase boundary may prove even more varied and exciting than that in either rotating bulk superfluid A or B phase: Consider, for example, Fig. 4, where the rotating phase boundary is illustrated for l adhered perpendicular to the A-B interface. With Ω parallel to l , the Cooper-pair angular momentum in $^3\text{He-A}$, an axisymmetric monopole structure forms on the $^3\text{He-B}$ vortex. However, upon reversal of Ω , a non-axisymmetric state necessarily follows. This may result as a bifurcation of the axisymmetric v vortex²⁶ into a half-quantum v -vortex molecule³⁰, just like in the observed vortex-core transition in rotating bulk superfluid $^3\text{He-B}$ (for reviews, see refs 6 and 7); such a structure occurs at distances of the order of (several) superfluid coherence lengths, $\xi_{GL} \approx 10(1 - T/T_c)^{-1/2}$ nm. The half-quantum pair may loop up into a vortex ring terminating at yet another monopole configuration.

The above ideas require experimental and further theoretical confirmation; our objective has been to discuss some of the exotic possibilities that will be encountered at the rotating superfluid A-B interface, and to motivate and encourage an experimental programme to verify such phenomena. We presently carry out extensive theoretical calculations including the above models; these will be reported in full elsewhere (M. M. Salomaa and G. E. Volovik, in preparation). Owing to the non-trivial connections to monopoles in quantum field theory^{8,31,32} and to domain walls and strings in cosmology³³, these investigations of the rotating A-B interface seem to afford interest in a general context, surpassing the physics of ultralow temperatures³⁴ and the superfluidity^{35,36} of ^3He .

I thank Professor Olli Lounasmaa, Dr. Yasumasa Takano, and the late Professor John Wheatley for conversations on experiments and Professors Gordon Baym, Keijo Kajantie, Anthony Leggett, and Lev Pitaevskii for theoretical conversations. I am grateful to Dr Gregory Volovik for enlightening discussions and useful comments on the manuscript. Thanks are also due to Professors Pradeep Kumar, Chia-Ren Hu, and Kazumi Maki for copies of ref. 14. This research is supported by the Academy of Finland.

Received 20 October; accepted 15 December 1986.

1. Zel'dovich, Ya. B., Kobzarev, I. Yu. & Okun, L. B. *Soviet Phys. JETP* **67**, 401-409 (1975).
2. Kibble, T. W. B. *J. Phys. A* **9**, 1387-1398 (1976).
3. Kibble, T. W. B. *Phys. Rep.* **67**, 183-199 (1980).
4. Vilenkin, A. *Phys. Rep.* **121**, 263-315 (1985).
5. Zurek, W. H. *Nature* **317**, 505-508 (1985).
6. Mineev, V. P., Salomaa, M. M. & Lounasmaa, O. V. *Nature* **324**, 333-340 (1986).
7. Salomaa, M. M. and Volovik, G. E. *Rev. mod. Phys.* (in the press).
8. Dirac, P. A. *Proc. R. Soc. A* **133**, 60-72 (1931).
9. Cabrera, B. *Phys. Rev. Lett.* **48**, 1378-1381 (1982).
10. Volovik, G. E. *JETP Lett.* **29**, 322-325 (1979).
11. Lavrentovich, O. D. *JETP Lett.* **43**, 382-385 (1986).
12. Volovik, G. E. & Mineev, V. P. *JETP Lett.* **23**, 593-595 (1976).
13. Blaha, S. *Phys. Rev. Lett.* **36**, 874-876 (1976).
14. Hu, C.-R., Kumar, P. & Maki, K. preprint, University of Southern California (1976).
15. Hall, H. E. *Physica* **90B**, 68-72 (1977).
16. Maki, K. *Physica* **90B**, 84-93 (1977).
17. Volovik, G. E. & Mineev, V. P. *Soviet Phys. JETP* **45**, 1186-1196 (1977).
18. Soni, V. *Phys. Rev. B* **18**, 6076-6081 (1978).
19. Bailin, D. & Love, A. *J. Phys. C* **11**, 1351-1359 (1978).
20. Wheatley, J. C. *Rev. mod. Phys.* **47**, 415-470 (1975).
21. Paulson, D. N., Krusius, M. & Wheatley, J. C. *J. Low Temp. Phys.* **25**, 699-715 (1976).
22. Buchanan, D. S., Swift, G. W. & Wheatley, J. C. *Phys. Rev. Lett.* **57**, 341-344 (1986).
23. Yip, S. & Leggett, A. J. *Phys. Rev. Lett.* **57**, 345-348 (1986).
24. Kaul, R. & Kleinert, H. *J. Low Temp. Phys.* **38**, 539-552 (1980).

25. Salomaa, M. M. *Phys. Rev. Lett.* (submitted).
26. Salomaa, M. M. & Volovik, G. E. *Phys. Rev. Lett.* **51**, 2040-2043 (1983).
27. Hindmarsh, M. & Kibble, T. W. B. *Phys. Rev. Lett.* **55**, 2398-2400 (1985).
28. Salomaa, M. M. & Volovik, G. E. *Phys. Rev. Lett.* **55**, 1184-1187 (1985).
29. Salomaa, M. M. & Volovik, G. E. *Phys. Rev. Lett.* **56**, 1313 (1986).
30. Salomaa, M. M. & Volovik, G. E. *Europhys. Lett.* **2**, 781-787 (1986).
31. 't Hooft, G. *Nucl. Phys. B* **79**, 276-284 (1974).
32. Polyakov, A. M. *JETP Lett.* **20**, 194-195 (1974).
33. Pagels, H. R. *Perfect Symmetry* (Simon & Schuster, Inc., New York, 1985).
34. McClintock P. V. E., Meredith, D. J. & Wigmore, J. K. *Matter at Low Temperatures* (Blackie & Son Ltd, Glasgow and London, 1984).
35. Volovik, G. E. *Soviet Phys. Usp.* **27**, 363-384 (1984).
36. Wölfle, P. & Vollhardt, D. *The Superfluid Phases of ^3He* (Taylor & Francis, London, in the press).

Organic polymer ferromagnet

Yu. V. Korshak*, T. V. Medvedeva*, A. A. Ovchinnikov† & V. N. Spector†

* Moscow Mendeleev's Institute of Chemical Technology, USSR

† Institute of Chemical Physics, USSR Academy of Sciences, Moscow, USSR

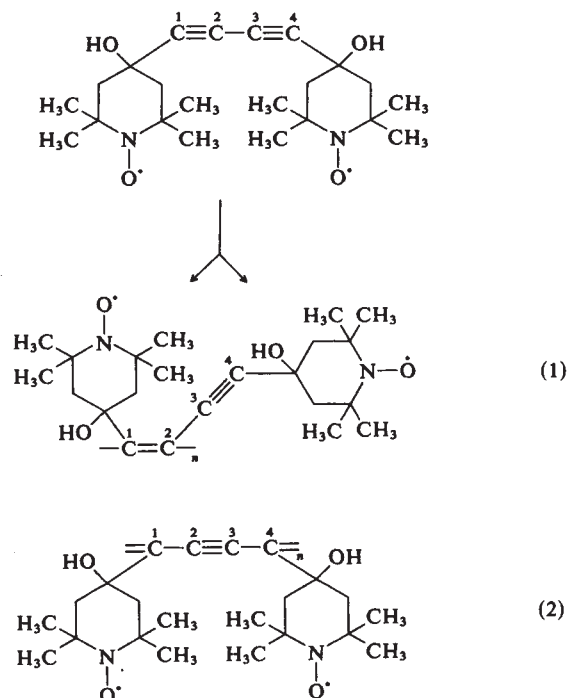
Low-dimensional, especially one-dimensional (1D), organic compounds are known to exhibit the properties of semiconductors, molecular metals and superconductors¹. The theoretical prediction of 1D organic ferromagnets and the principles of construction of polymer polyradicals with a ground state spin proportional to a number of monomeric units were proposed in 1977-78^{2,3}. For nearly two decades various hydrocarbons of high spin multiplicity have been studied as the model of an organic ferromagnet⁴⁻⁶. Advanced quantum chemistry techniques have been applied for theoretical investigations of the electronic structure of such compounds⁷⁻⁹. Here we present a study of an organic polymer ferromagnetic preparation with a transition metal impurity content well below the level likely to significantly affect its magnetic properties. We find that all the organic polymer ferromagnet samples can be conditionally subdivided into three groups: paramagnetic polymers⁷⁻⁹, spin glasses¹⁰ and real polymer ferromagnets.

An organic ferromagnet was prepared by polymerization of the paramagnetic stable biradical, 1,4-bis-(2,2,6,6-tetramethyl-4-oxy-4-piperidyl-1-oxyl)-butadiin, (BIPO)¹¹. The transition metal content was controlled by methods including atomic absorptional analysis and laser amplified mass-spectrometry, with at least two samples of the preparation used for each run. The samples with maximum concentration of iron, C(Fe), $\sim 6.5 \times 10^{-5}$ and total concentration of the transition metals, (Fe, Co, Ni and so on), C(tr.m.tot) $\leq 1 \times 10^{-4}$ mass% were used for magnetic measurements. It was noted, that both C(Fe) and C(tr.m.tot) were $\sim 2-3$ times smaller in the polymer (polyBIPO) than in the initial monomer. An independent analysis of the separated magnetic and non-magnetic fractions of the polymer preparation by K. Kimura from the Institute for Molecular Science, Japan indicated that the possible concentration of the transition metals was below any significant level.

Biradical monomers form perfect needle monocrystals. An X-ray analysis by V. Shklower *et al.* (Fig. 1) gives the symmetry group Pccn, $Z=8$, with the following cell parameters: $a=19.2425$; $b=16.2402$; $c=14.1883$ Å; $V=4,433$ Å³, and the esti-

mated X-ray density $d=1.17$ g cm⁻³. Thermal dependence of the monomer magnetic susceptibility is described well by the formula: $X_m(T)=C/(T+\theta_a)$ with the constants $C=7.7 \times 10^{-4}$ grad cm³ g⁻¹ and $\theta_a \sim 2$ K. Such susceptibility dependence indicates antiferromagnetic ordering at about 1 K. T_N is found by extrapolation. As in the present paper all the magnetic measurements have been taken with the Oxford Instruments model of Faraday balances, allowing temperatures >1.7 K¹¹.

Biradical monomers undergo polymerization under the action of a number of factors, such as heating, ultraviolet irradiation, a glowing discharge, and so on. The kinetics of thermal polymerization were studied photometrically using the absorption bands at 293 and 325 nm (conjugated C=C bonds), arising in addition to the bands at 204 and 242 nm in ultraviolet spectra of monomer preparations. An autocatalytic acceleration of the reaction was observed at 80 °C, starting at approximately 20% conversion, which had also been noted for a number of other diacetylenes¹². With the rise of temperature up to 90-100 °C the induction period disappeared and the conversion curves exhibited typically asymptotic character. An increase of the molecular mobility in the monomer monocrystal may be a possible cause for the transition to a non-inductional process. The X-ray data for monomers indicate the impossibility of topochemical polymerization. Nevertheless polymerization involves up to 100% conversion and an explosive development of the process (with conversion $\sim 80\%$) has been noted in some cases. Powder patterns of the polymer samples¹³ display an absence of far orders and a high stacking anisotropy (>0.3), although the outer form of the initial monomer crystals may stay unaltered even at 100% conversion to polymer. These facts are consistent with the two possible reaction paths:



Here pathway (1) is a process 1,2 addition and pathway (2) is a 1,4 non-topochemical addition.

Chemical analysis data (Table 1) for monomer and polymer samples indicate the preservation of the chemical composition and the absence of secondary reactions. Similar results are obtained for photo-stimulated polymerization and polymerization in a glowing discharge.

The two proposed reaction pathways are also consistent with the results of infrared spectroscopy for BIPO and polyBIPO. The band in the region 1,600-1,700 cm⁻¹ in infrared spectra (with a resolution of 4 cm⁻¹) could be produced by a polymer

Table 1 Chemical analysis data for monomer (BIPO) and polymer (polyBIPO) samples

Element	Calculated BIPO	mass % BIPO	Experimental polyBIPO
C	67.66	67.30	67.90
H	8.77	9.02	9.00
N	7.17	7.16	7.09
Transitional metals	—	$\leq 2 \times 10^{-4}$	$\sim 6.5 \times 10^{-5}$

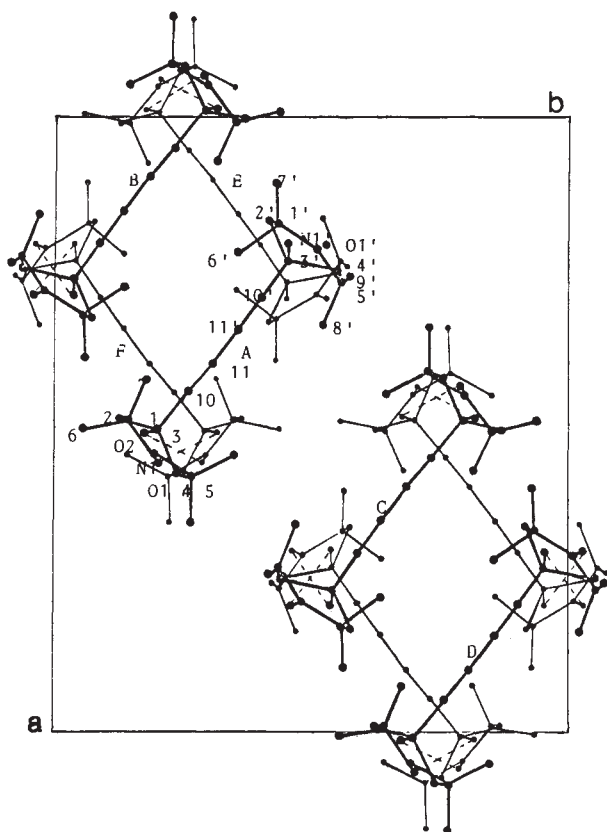


Fig. 1 The *ab* projection of the crystal structure of BIPO. The basis molecule A and molecule B, related to the former by the $2(1/4, 1/4z)$ axis and situated at the same height along *z*-axis; molecule C, related to basis one by the symmetry centre $\bar{1}(1/2, 1/2, 1/2)$, and molecule D, related to C molecule by the $2(3/4, 3/4, z)$ axis, form an infinite gittered tape along $[112]$ direction with the folds at every diacetheline rod (dar) of 153.8° . The molecules A, B, C, and D belonging to the same tape, are shown by the bolder lines. The distance between the gravity centres of dar for A and B molecules of the same tape is $5.47 \text{ \AA}$ and for A and C molecules is $8.06 \text{ \AA}$. The vector, connecting the gravity centres of dar A and B, forms with the rod of A molecule an angle 71° . H-bonds of the type $O(1)-H[O(1)] \cdots O(2')$ of 2.831 and $O(2)-H[O(2)] \cdots O(1')$ of $2.768 \text{ \AA}$ (dotted lines) array the basis molecule (A) with two pairs of E and F molecules, related to A molecule by the glide planes parallel to *ac* and *bc* faces, into infinite H-bonded associates parallel to *c* axes. The distance between the planes, including dar A and B, and the corresponding planes, including the molecules E and F (situated both over and below them), is $c/2 = 7.094 \text{ \AA}$. The molecules E and F belong to the other family of the gittered tapes, related to the former ones by 2_1 axes parallel to *a* axis and extended along $[1\bar{1}2]$.

of conjugated double bonds ($C=C$). This does not contradict the existence of either structure (1) or structure (2). The band with a maximum at $2,146 \text{ cm}^{-1}$ may be associated with a presence of $C\equiv C$ bonds in the main chain of a macromolecule (if it is not due to overtones). Some finite concentration of the carbene centres entangled in a solid polymer matrix follows from electron spin resonance (ESR) spectra by A. Tzapin ($g = 4.00\text{--}4.15$, manuscript in preparation). This does not rule out a polymerization process as the reaction of 1,4-non-topochemical addition (2).

The conformational analysis¹⁴ of BIPO and polyBIPO in vitreous solvent (C_2H_5OH at 77 K) has shown the decrease in distance between the nitroxyl fragments in the polymer in comparison with the monomer, which may be one of the physical reasons for the strengthening of an exchange interaction between these fragments. The concentration of the stable radicals as measured by A. Shapiro remains essentially constant during

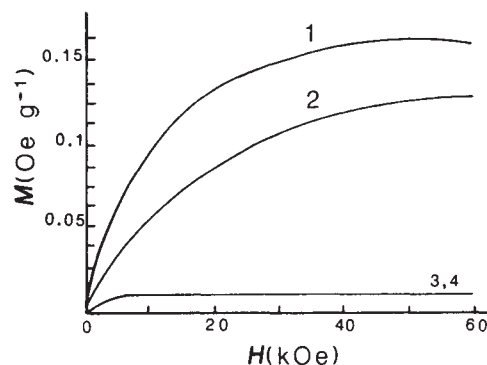


Fig. 2 Field dependence of polyBIPO magnetization at 1.7, 4.2, 25 and 90 K (curves 1–4, correspondingly).

polymerization. The total fall in nitroxyl concentration at 100°C does not exceed 1.0–1.5% at 75% conversion to polymer.

The polymerization degree and molecular mass distribution of polymer preparations vary very widely whatever kind of stimulation of monomer conversion is used. This variation leads to a variation of the magnetic properties of polymer preparations according to theoretical predictions^{2,3}. In polymer samples displaying spontaneous magnetization, we obtained values ranging from 10^{-2} to >1 . This report is mainly concerned with polymers prepared by the photo-stimulated monomer conversion.

When exposed to light a monomer orange sample transforms into a black powder as polymerization is completed. We studied the magnetization via field dependencies of these polymer samples over a broad temperature range, and our findings at 1.7, 4.2, 25 and 90 K are shown in Fig. 2. The diamagnetic component has been subtracted from the sample magnetic moment value at 25 and 90 K. At low temperatures (1.7 and 4.2 K) the paramagnetic component, being of the order of $0.5 \times 10^{19} \text{ spin cm}^{-3}$, is dominant and so only curves 3 and 4 are illustrative of spontaneous magnetization and are typical for ferromagnets. The measured value of the spontaneous magnetization is only 0.1% of its theoretical value for 1D organic ferromagnets^{2,3}, which implies that only an insignificant fraction of the particles possess ferromagnetic properties. The fact that the magnetization value changes from one sample to another may be explained by the growth conditions (as yet practically uncontrollable) of the superstructural particles (and consequently magnetic domains) in the course of the polymerization process. Nevertheless magnetic separation allows the selection of particles with a magnetization value above 1 G, which respond noticeably to the field of a permanent magnet. The results of the magnetic measurements of the enriched polymer samples will be reported later.

ESR spectra ($\nu = 3 \text{ cm}^{-1}$) show two broad bands with maxima at $H_1 = 1,800$ and $H_2 = 5,000 \text{ Oe}$ reproducible for different samples. They correspond to ferromagnetic resonance in ferromagnets of an easy-axis type with an anisotropy field $H_A \sim 2,400 \text{ Oe}$. It is to be noted that the demagnetizing factors and the porosity of a sample were ignored in the analysis of the curves (Fig. 3)¹, obtained by A. Shapiro. The field $H_A = 2,400 \text{ Oe}$ corresponds well to the zeroth fields in stable organic biradicals and in organic molecules in a triplet state¹⁵. PolyBIPO samples demonstrate an expressed magnetic anisotropy (for example, curves 1 and 2 in Fig. 4), the reason for which (from the point of view of a spin-glass approach and more detailed experiments with A. Tzapin) will also be reported later.

Measurements of the Curie temperature in the range $0\text{--}310^\circ \text{C}$ with a vibrational magnetometer (field $= 0.04 \text{ T}$) showed that generally the magnetic properties of polyBIPO were preserved. For some samples the Curie temperature was $150\text{--}190^\circ \text{C}$ ¹¹. For the samples of the higher magnetization the presence of ferromagnetism was observed up to 310°C (fast decomposition). Such noticeable differences in the Curie temperatures may be

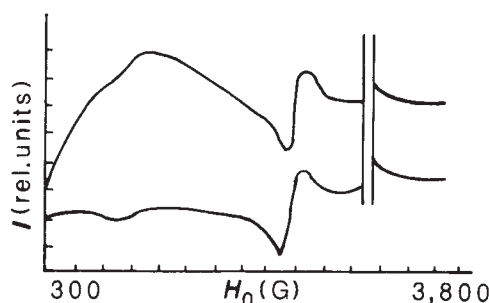


Fig. 3 ESR spectra of polyBIPO in two arrangements [11]. 1, Parallel to the field; 2, perpendicular to the field.

connected with the specific morphological features of the different samples and consequent variations of the degradation and destruction processes at elevated temperatures. A temperature of the same order of magnitude could be found by noting that polyBIPO is a quasi-1D, highly anisotropic Heisenberg ferromagnet. The scale value of the transverse exchange integrals can be found from the value of θ_a for the monomer ($J_{\perp} - 2 \text{ K}^{11}$). The exchange integrals along a chain are $J_{\parallel} = 1 \text{ eV}$. Such a large value of $J_{\parallel}$ is connected with the phenomenon of spin transfer from the radical to the main chain (in preparation).

Received 9 April; accepted 9 December 1986.

1. Miller, J. *Extended Linear Chain Compounds* Vols 1-3 (Plenum, New York, 1983).
2. Ovchinnikov, A. A. *Dokl. (Proc.) Acad. Sci. USSR* **236**, 928-931 (1977).
3. Ovchinnikov, A. A. *Theor. chim. Acta* **47**, 297-303 (1978).
4. Ovchinnikov, A. A. & Cheranovskii, V. O. *Teor. éksp. Khim.* **17**, 1010-1018 (1981).
5. Tyutyulkov, N. N., Polansky, O. E., Schuster, P., Karabunarliev, S. & Ivanov, C. J. *Theor. chim. Acta*, **67**, 211-228 (1985).
6. Nasu, K. *Phys. Rev.* **B33**, 330-338 (1986).
7. Jtoh, K. *Chem. Phys. Lett.* **1**, 235-238 (1967).
8. Jtoh, K. *Pure appl. Chem.* **50**, 1251-1259 (1978).
9. Teki, Y., Takui, T., Jtoh, K., Jwamura, H., Kobayashi, K. *J. Am. chem. Soc.* **108**, 2147-2156 (1986).
10. Korshak, Yu. V., Ovchinnikov, A. A., Shapiro, A. M., Medvedeva, T. V. & Spector, V. N. *Zh. éksp. teor. Fiz. Ptiima*, **43**, 309-311 (1986).
11. Baughman, R. H. *J. Polym. Sci.* **12**, 1511-1536 (1974).
12. Korshak, Yu. V., Ovchinnikov, A. A., Shapiro, A. M., Medvedeva, T. V. & Spector, V. N. *Abstr. int. Conf. Sci. Technol. synthetic Metals*, Kyoto (1986).
13. Bloor, D., Ando, D. I., Fisher, D. A. & Hubble, C. L. in *Molecular Metals* (ed. Hatfield, W. E.) 249 (Plenum, New York, 1979).
14. Kerrington, A. & McLachlan, K. A. *Magnetic Resonance and its Chemical Applications* (Harper & Row, New York and Evanstone, London, 1967).

Icosahedral symmetry versus local icosahedral environments in Al-Mn alloys from NMR

L. H. Bennett*, J. W. Cahn*, R. J. Schaefer*,
M. Rubinstein† & G. H. Stauss†

* National Bureau of Standards, Gaithersburg, Maryland 20899, USA

† Naval Research Laboratory, Washington DC 20375, USA

In a recent report¹, Pauling suggested that the icosahedral diffraction found in the aluminium-manganese system is due to a cubic crystalline phase with a large unit cell of more than 1,000 atoms, multiply twinned to mimic icosahedral symmetry, instead of a crystalline or quasicrystalline phase with fivefold symmetry. Pauling claims that most of the manganese atoms are surrounded by twelve aluminium atoms arrayed in a nearly perfect icosahedron, a cluster that occurs in the Al_{12}W structure. To account for the high Mn content in the icosahedral phase, he assumed that each aluminium atom was shared with a neighbouring icosahedron. Here, we report a comparison of the nuclear magnetic resonance (NMR) spectra obtained from aluminium-manganese alloys in

the crystalline 'G' phase (in which each of the Mn atoms is surrounded by an icosahedron of Al atoms, exactly as in Pauling's proposal structure) with the spectra of the icosahedral phase. We find that the NMR spectrum in the 'G' phase is entirely different from the spectrum in the icosahedral phase. Twinning should not affect the NMR spectra except for a small contribution from atoms on or near the composition planes, so the Al-Mn icosahedral phase is not Pauling's twinned cubic structure.

No known phase of aluminium and manganese has icosahedra bonded in the way Pauling suggests. However the 'G' phase of Al-Mn, having the Al_{12}W structure, was reported² to exist as a metastable precipitate. In the G phase, the Mn-centred Al icosahedra are arranged on a body-centred cubic (b.c.c.) lattice, without sharing of Al atoms by neighbouring icosahedra. We have reinvestigated this G phase³ and found that it becomes stable at temperatures $< 500^\circ\text{C}$.

To form the G phase sample, a liquid aluminium alloy with 6.3 at.% Mn was first rapidly solidified by melt spinning to give a mixture of icosahedral phase and supersaturated face-centred cubic (f.c.c.) aluminium. During subsequent heating at 400°C , the melt-spun material transforms within a few minutes to a fine-grained mixture of orthorhombic Al_6Mn crystalline phase plus saturated f.c.c. aluminium. This mixture is stable above 520°C , but metastable below 500°C . The G phase forms in about 1,000 h at 400°C by the slow reaction of Al_6Mn with Al, creating material that is more than 80% Al_{12}W structure, with the balance being f.c.c. aluminium containing less than 1% Mn in solid solution.

To form the icosahedral phase sample, a liquid aluminium alloy with 15 at.% Mn was rapidly solidified by melt spinning without subsequent heat treatment. The alloy contained about 10% f.c.c. aluminium.

We have already reported⁴ on the ^{27}Al and ^{55}Mn NMR powder pattern lineshapes of icosahedral Al-Mn. Both the Al and Mn resonances indicate broad distributions of local environments⁴. (This result was also reported by Warren *et al.*⁵) The ^{55}Mn NMR of the icosahedral phase is shown in Fig. 1, and compared with the ^{55}Mn NMR in the G phase. From the structure, it is known that there is a single local Mn environment in the G phase with zero quadrupole interaction, as indicated by its extremely narrow peak. In contrast, the ^{55}Mn peak of the icosahedral phase is broad with no detectable sites (that is, $< 5\%$ of the intensity) with zero or near zero quadrupole interaction.

Atoms sitting on sites of either cubic or icosahedral symmetry show zero quadrupole interactions. The Mn atoms in the G phase occupy sites of cubic symmetry, their nearest neighbours being 12 Al atoms arranged in the form of a nearly perfect icosahedron. As a result, the ^{55}Mn NMR spectrum of the G phase shows a zero quadrupole interaction. In Pauling's model of the icosahedral phase, the majority of the Mn atoms have a nearest-neighbour environment identical to that in the G phase except for a distortion of the icosahedron, not specified, but presumed to be minor. Although most of the Mn atoms do not sit on positions of exact cubic or icosahedral symmetry in Pauling's model, such a structure should produce a ^{55}Mn NMR spectrum only slightly broadened from that of the G phase. (A three per cent stretching of an icosahedron produces a quadrupole interaction less than ten per cent of the quadrupole interaction from a single uncompensated charge.) We have found, however, that the ^{55}Mn NMR spectrum of the icosahedral phase is completely different from that of the G phase: the icosahedral phase contains no detectable sites with zero or near zero quadrupole interaction, showing that the Mn atoms do not lie at sites that are within the range of distortions of symmetric nearest neighbour environments expected with Pauling's model. The specific icosahedral cluster is a central part of this model as it led him to accurate agreement with distances, but not directions, in the diffraction data. The NMR data thus rule out both Pauling's cluster and structural model for the Al-Mn icosahedral phase.

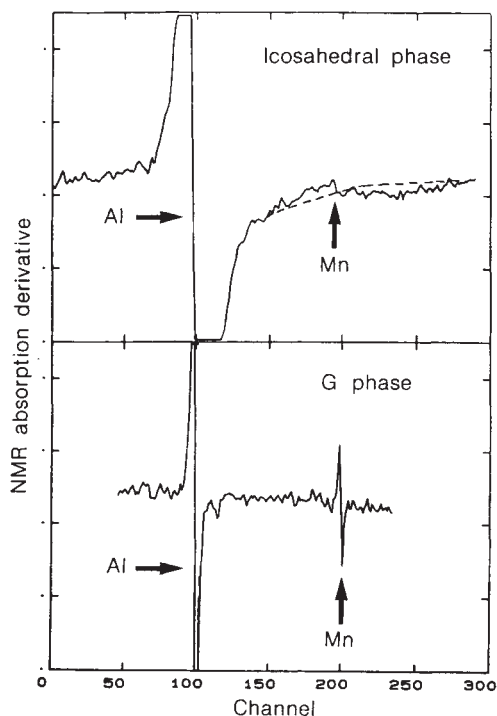


Fig. 1 Comparison of the ^{55}Mn and ^{27}Al nuclear magnetic resonance derivative powder pattern taken at 11.037 MHz at room temperature in icosahedral Al-Mn and in the G phase, Al_{12}Mn . The arrows point to the centre of the resonances. In the iAl-Mn, the ^{55}Mn resonance lies in the wings of the ^{27}Al NMR. The Mn resonance in iAl-Mn is revealed by subtracting the continuation of the Al resonance shown by the dashed curve. The ^{27}Al NMR in the G phase is largely masked by the free Al component, but in any case is considerably narrower than the corresponding spectrum for the icosahedral phase. Each channel number is equal to 5.62 Oe.

It is important to distinguish between the symmetry of assemblies of atoms such as crystals or quasicrystals as detected by diffractions, and symmetries of sites occupied by atoms in such assemblies, as detected by probes such as NMR^{4,5} and the Mossbauer effect (NGR)^{6,7}. (The NMR spectra confirm Mossbauer effect results obtained on the icosahedral phase with some iron atoms substituted for manganese, in which it was assumed that the iron atoms would occupy the same sites as the manganese.) High-symmetry crystals or quasicrystals can occur without any atom occupying a high-symmetry position. For example, the $\alpha\text{Al-Mn-Si}$ cubic phase contains icosahedral clusters, with a vacancy at the centre of each. No atom occupies a position of local cubic symmetry. Conversely, highly symmetric clusters often pack into crystals of low symmetry. It is therefore not surprising that the Mn atoms in the Al-Mn icosahedral phase do not occupy positions of high local symmetry. Indeed their local environment seems to be⁹ quite similar to that of Mn in the $\alpha\text{Al-Mn-Si}$ structure. In a number of recently proposed models⁷⁻¹⁴ for the icosahedral phase, the Mn atoms are packed in several ways to produce a structure with long-range icosahedral symmetry, but the nearest-neighbour environment of the Mn atoms is highly unsymmetric. The measured NMR spectrum of the icosahedral phase is consistent with such models. Any model of the Al-Mn icosahedral phase must provide diffraction displaying icosahedral symmetry, but it must also have complex local environments as revealed by the experiments.

Note added in proof: L. Pauling, in commenting on the present work, points out that only 8 of the 195 Mn atoms per cell in his structure lie in positions with cubic point-group symmetry and thus have zero quadrupole broadening. The quadrupole interactions of the other icosahedrally coordinated Mn atoms must be

non-zero, but we can estimate that they would be too small to replicate the NMR spectrum of the icosahedral phase. Although the exact equivalence of Pauling's structure based on the packing of Al icosahedra with Mn at their centres to the NaCd_2 structure is not clear to us, we have modelled the quadrupole interaction at every site in the NaCd_2 structure, and in no case do we find values comparable to those observed. Further distortions of the structure, large enough to fit the data, would render the conceptual representation of the structure by icosahedra invalid.

Received 28 August 1986; accepted 16 January 1987.

1. Pauling, L. *Nature* **317**, 512-514 (1985).
2. Little, K., Raynor, G. V. & Hume-Rothery, W. *J. Inst. Metals* **73**, 83-90 (1946).
3. Schaefer, R. J., Biancaniello, F. S. & Cahn, J. W. *Scr. met.* **20**, 1439-1444 (1986).
4. Rubinstein, M., Stauss, G. H., Phillips, T. E., Moorjani, K. & Bennett, L. H. *J. mater. Res.* **1**, 243-246 (1986).
5. Warren, W. W., Jr, Chen, H. S. & Hauser, J. J. *Phys. Rev. B* **32**, 7614-7616 (1985).
6. Swartzendruber, L. J., Shechtman, D., Bendersky, L. & Cahn, J. W. *Phys. Rev. B* **32**, 1383-1385 (1985).
7. Eibschutz, M., Chen, H. S. & Hauser, J. J. *Phys. Rev. Lett.* **56**, 169-172 (1986).
8. Elser, V. & Henley, C. L. *Phys. Rev. Lett.* **55**, 2883-2886 (1985).
9. Stern, E. A., Ma, Y. & Bouldin, C. E. *Phys. Rev. Lett.* **55**, 2172-2175 (1985).
10. Mackay, A. *Acta crystallogr.* **15**, 916-918 (1962).
11. Audier, M. & Guyot, P. *Phil. Mag.* **53**, L43-L51 (1986).
12. Kuriyama, M., Long, G. G. & Bendersky, L. *Phys. Rev. Lett.* **55**, 849-851 (1985).
13. Watson, R. E. & Weinert, M. *Mater. Sci. Eng.* **79**, 105-109 (1986).
14. Cahn, J. W. & Gratias, D. *J. Phys. Colloq.* **C3**, 415-424 (1986).

Satellite measurements of sea surface cooling during hurricane Gloria

Peter Cornillon*, Lothar Stramma† & James F. Price‡

* Graduate School of Oceanography, University of Rhode Island, Kingston, Rhode Island 02881, USA

† Institut für Meereskunde, Kiel 2300, FRG

‡ Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

Hurricanes and other strong storms can cause important decreases in sea surface temperature by means of vertical mixing within the upper ocean, and by air-sea heat exchange. Here we use satellite-derived infrared images of the western North Atlantic to study sea surface cooling caused by hurricane Gloria (1985). Significant regional variations in sea surface cooling are well correlated with hydrographic conditions. The greatest cooling (up to 5 °C) occurred in slope waters north of the Gulf Stream where the seasonal thermocline is shallowest and most compressed; moderate cooling (up to 3 °C) occurred in the open Sargasso Sea where the thermocline is deeper and more diffused; little or no cooling occurred in shallow coastal waters (bottom depth less than 20 m) which were isothermal before the passage of hurricane Gloria. There is a pronounced right-side asymmetry of sea surface cooling with stronger (by a factor of 4) and more extensive (by a factor of 3) cooling found on the right side of the hurricane track. These qualitative results are consistent with the notion that vertical mixing within the upper ocean is the dominant sea surface cooling mechanism of hurricanes.

Hurricane Gloria (1985) began to form very late in the hurricane season off the Cape Verde Islands¹. Gloria moved nearly due westward with the trade winds to about the Leeward Islands, and then turned north-west toward the Sargasso Sea. On 25 September the minimum central pressure fell to 919 mbar, making Gloria one of the most powerful Atlantic hurricanes of the past decade. During 26 and 27 September, the central pressure slowly increased as Gloria moved northward at about 7 ms⁻¹ over the Sargasso Sea and coastal waters of the north-east United States^{1,2} (Fig. 1). Gloria crossed the US coast over Long Island on the afternoon of 27 September. Although considerably weakened by then, Gloria caused substantial property damage in southern New England².

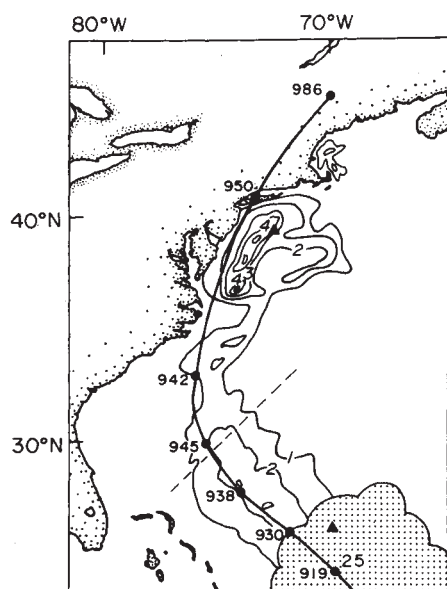


Fig. 1 Magnitude of sea surface cooling observed from satellite infrared images. Contour intervals are 1°C . The thicker line is the path of hurricane Gloria, which moved to the north; dots are 12 h apart. The number on the right of the southernmost dot along the track refers to the day in September 1985 (at 00:00 GMT). The numbers on the left of the track are hurricane central pressure in mbar. Solid triangles show the location of two expendable bathythermographs (XBTs) that provided the data of Fig. 3, and the dashed diagonal line crossing 30°N , 75°W is the SST section of Fig. 4. Stippled area at lower right is clouds.

To see what effects Gloria had on the ocean, we have examined satellite infrared images of the western North Atlantic made by the NOAA-9 Advanced Very High Resolution Radiometer (AVHRR)³. The data were decimated to give 4-km resolution, corrected for atmospheric attenuation⁴, and scan-angle corrected⁵. Previous studies⁴ have shown that the resulting sea surface temperature (SST) data have an absolute accuracy of about 0.5°C , and a relative error (point-to-point variation) of about 0.2°C .

A pre-hurricane SST image was formed by compositing night-time images of the western North Atlantic made during 19–23 September (Fig. 2, left). The compositing procedure selected the warmest pixel from the sample in order to reduce the area obscured by cloud cover. (Night-time images were chosen in order to avoid contamination by diurnal warming which can be strong in this region^{4,5}.) A post-hurricane composite image was formed for the period immediately following the hurricane passage, 27 and 28 September, and by also ensuring that only those regions south of (behind) the hurricane were retained (Fig. 2, right). The post-hurricane data used here are thus less than two days old and nearly synoptic, which helps minimize the unwanted effects of relaxation⁶ and distortion by horizontal advection (though advection by the Gulf Stream is apparent as noted below). An estimate of the change in SST caused by the passage of Gloria was obtained by differencing the pre- and post-hurricane SST images (Fig. 2, centre). The same data were also smoothed and contoured to yield Fig. 1. These images are unique, we believe, in showing such a large portion of a hurricane track and in covering a region having a wide range of hydrographic conditions.

The pattern of sea surface cooling appears to be dominated by the response to hurricane Gloria except near the Gulf Stream axis where there is a clear distortion due to horizontal advection. Note the northeastward-extending notch in the cooling contours at the intersection of the hurricane track and the Gulf Stream axis (near 35°N), and the apparent northeastward displacement of cooled water by as much as 150 km downstream of the Gulf Stream-track intersection (Fig. 1 and Fig. 2, centre).

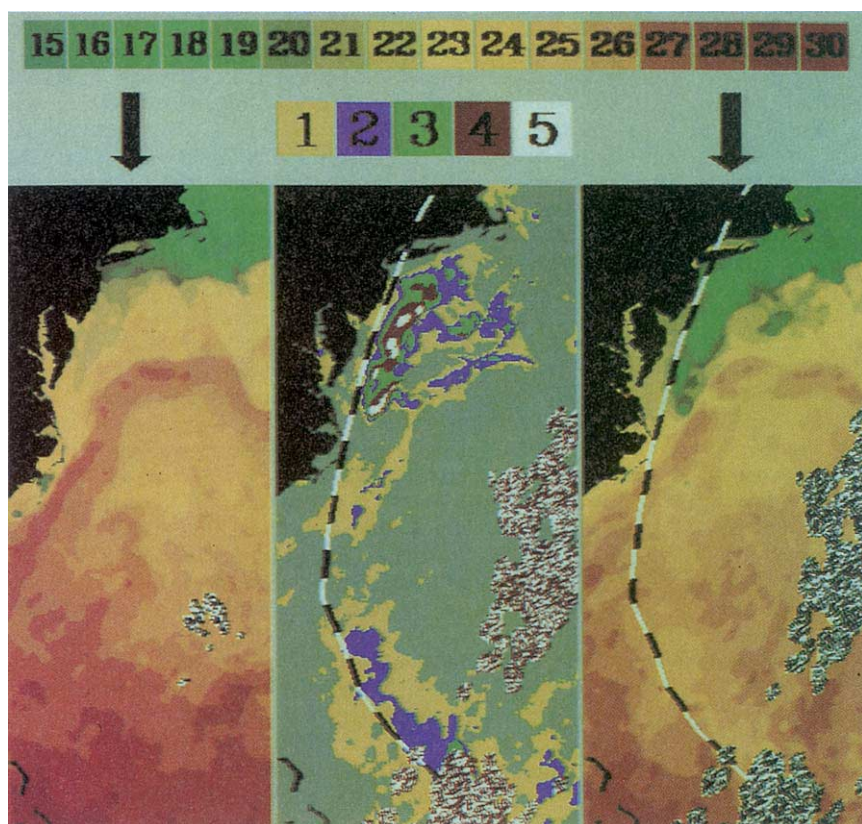


Fig. 2 Composite of pre-hurricane night-time satellite images of the western North Atlantic for 19–23 September 1985 (left). Post-hurricane composite for 27 and 28 September (right). Difference between the pre- and post-hurricane images (centre). Black, masked area is the east coast of the United States from Cape Hatteras to Cape Cod, and stippled areas denote cloud cover. The colour scale at the top of the figure is temperature scale (in $^{\circ}\text{C}$) for the left and right panels. The second colour scale from the top (range $1\text{--}5^{\circ}\text{C}$) applies to the central panel.

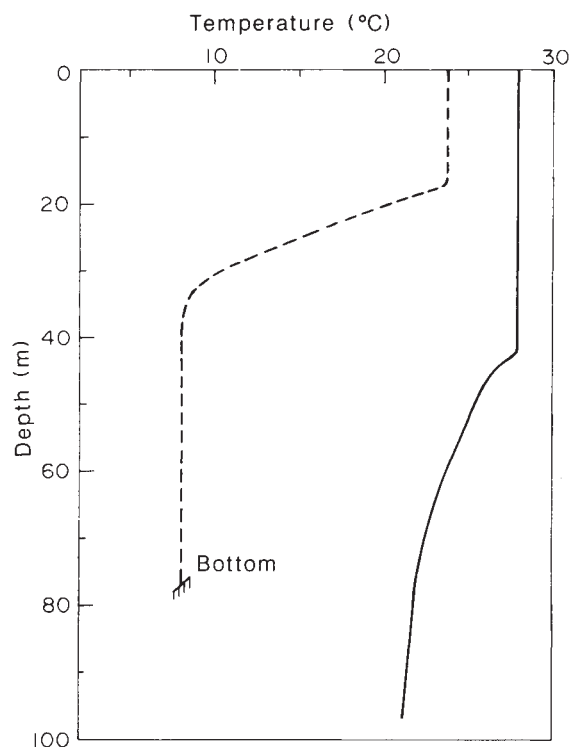


Fig. 3 Pre-hurricane, XBT-derived temperature profiles from slope water north of the Gulf Stream (dashed line, 39°N, 72°W, 11 September) and from the southern Sargasso Sea (solid line, 26°N, 70°W, 18 September) (see Fig. 1 for locations).

These images show a significant regional (or along-track) variation of sea surface cooling which gives useful clues to the mechanism of cooling. Previous model studies have shown that vertical mixing of cooler thermocline waters up into the ocean's surface layer is the dominant mechanism producing sea surface cooling^{7,8,9}, and that sea surface cooling by a given storm will be greatest where cold thermocline waters are closest to the sea surface. These images support this result as far as regional variations of sea surface cooling appear correlated with hydrographic conditions rather than with hurricane intensity or initial SST. Most notably, hurricane intensity (measured by central pressure, Fig. 1) decreases almost monotonically from south to north along the portion of the hurricane track seen here^{1,2}. On the other hand, the strongest sea surface cooling, up to 5°C, is observed in slope waters to the north of the Gulf Stream where Gloria had already weakened. Along Gloria's track, cold thermocline water is found nearest the sea surface in slope waters where the thermocline is shallow and very compressed compared to the thermocline in the Sargasso Sea (Fig. 3). Thus, hurricane-induced vertical mixing would be expected to reduce the SST more effectively over slope waters than over the Sargasso Sea. However, over shallow coastal waters (bottom depth <20 m; for example, Long Island Sound, Rhode Island Sound, and Nantucket Sound) there was little or no sea surface cooling. By late September these shallow waters are usually already isothermal¹⁰ (Fig. 3), so that further vertical mixing by Gloria could not produce further cooling.

Sea surface cooling shows a striking asymmetry about the track (Fig. 1), with much stronger and more extensive decreases in SST found on the right side of the track compared to the left side (viewing the sea surface from above, and assuming that the hurricane moves toward the top of the figure). This right-side asymmetry has been observed before in conventional, *in situ* data sets^{7,11} and in other satellite images⁶, though not as clearly. In this case, some of the asymmetry found within coastal waters is attributable to the horizontal variations in hydrographic condi-

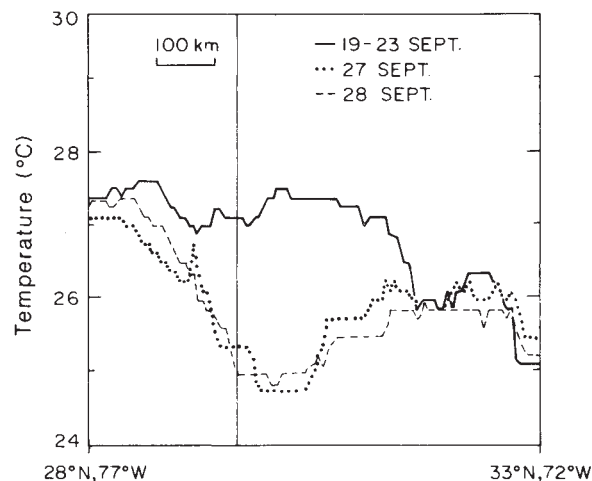


Fig. 4 Satellite-derived 5×5 pixel median-filtered SST sections between 28°N, 77°W and 33°N, 72°W (dashed diagonal line of Fig. 1). A section made through the pre-hurricane composited image is shown by the solid line (19–23 September). Sections made from post-hurricane images are shown by a dotted line (27 September, 07:57 GMT) and by a dashed line (28 September, 07:46 GMT). The thin vertical line to the left of centre marks the intersection with Gloria's track.

tions noted above. Over the open Sargasso Sea this right-side asymmetry of cooling is thought to be a characteristic of the homogeneous ocean's response to a hurricane. (Except in the immediate vicinity of the Gulf Stream, the near-surface currents are on average westward^{12,13}, so that advection is unlikely to be the cause of the asymmetry.) To quantify the cooling asymmetry, we have sampled the pre- and post-hurricane SST images along a section across the Sargasso Sea (Figs 1 and 4). The maximum sea surface cooling, about 2.6°C, occurs at distance $R \approx 80$ km to the right of the track. At the same distance to the left of the track, the sea surface cooling is only about 0.6°C. The ratio $2.6/0.6 \approx 4$ provides one measure of the sea surface cooling asymmetry along this section, and by inspection of Fig. 1, a value characteristic of the cooling throughout the Sargasso Sea. As an alternative measure of asymmetry, we have also calculated the line integral of sea surface cooling from the track out to distances of $4R$ to the right and to the left of the track. The integral cooling to the right of the track is about 550°C km, and to the left of the track the value is about 180°C km. These give a ratio of roughly 3. A similar, strong right-side asymmetry of sea surface cooling occurs in numerical models which presume that vertical mixing in the upper ocean is mainly a consequence of shear flow instability of wind-driven currents^{7,8}.

We thank O. Brown, R. Evans, J. Brown and A. Li at the University of Miami for their assistance with image processing software; D. Evans and R. S. Armstrong for the XBT profiles; and J. Rahn for editorial comments. Supported by contracts N00014-81-0062 (P.C. and L.S.) and N00014-84-C-0134 (J.F.P.) with the US Office of Naval Research.

Received 14 July 1986; accepted 13 January 1987.

- Case, R. A. & Gerrish, H. P. *Mariners Weath. Log* **30**, 6–11 (1986).
- DeAngelis, D. *Mariners Weath. Log* **30**, 34–41 (1986).
- Schwalb, A., *Natn. Oceanogr. Atmos. Admin. tech. Memo NESS 95* (Washington DC, 1978).
- Stramma, L., Cornillon, P., Weller, R. A., Price, J. F. & Briscoe, M. G. *J. phys. Oceanogr.* **16**, 827–837 (1986).
- Cornillon, P. & Stramma, L. *J. geophys. Res.* **90**, 11811–11815 (1985).
- Stramma, L., Cornillon, P. & Price, J. F. *J. geophys. Res.* **91**(C4), 5031–5035 (1986).
- Price, J. F. *J. phys. Oceanogr.* **11**, 153–175 (1981).
- Martin, P. J. *J. geophys. Res.* **87**, 409–427 (1982).
- Chang, S. W. & Anthes, R. A. *J. phys. Oceanogr.* **8**, 468–480 (1978).
- Riley, G. A. *Bull. Bingham oceanogr. Coll.* **15**, 15–46 (1956).
- Pudov, V. D., Varfolomeyev, A. A. & Fedorov, K. N. *Oceanology* (English transl.) **18**, 142–146 (1978).
- Beardsley, R. C. & Boicourt, W. C. *Evolution of Physical Oceanography 198–233* (MIT Press, Cambridge, 1981).
- Worthington, L. V. *Johns Hopkins oceanogr. Stud.* **6** (1976).

A mechanism for the west-north-west movement of monsoon depressions

B. N. Goswami

Centre for Atmospheric Sciences, Indian Institute of Science, Bangalore 560 012, India

The monsoon depressions intensify over the Bay of Bengal, move in a west-north-west (WNW) direction and dissipate over the Indian continent. No convincing physical explanation for their observed movement has so far been arrived at, but here, I suggest why the maximum precipitation occurs in the western sector of the depression and propose a feedback mechanism for the WNW movement of the depressions. We assume that a heat source is created over the Bay of Bengal due to organization of cumulus convection by the initial instability. In a linear sense, heating at this latitude (20°N), produces an atmospheric response mainly in the form of a stationary Rossby-gravity wave to the west of the heat source. The low-level vorticity (hence the frictional convergence) and the vertical velocity associated with the steady-state response is such that the maximum moisture convergence (and precipitation) is expected to occur in the WNW sector at a later time. Thus, the heat source moves to the WNW sector at a later time and the feedback continues resulting in the WNW movement of the depressions.

Monsoon depressions are one of the most important synoptic-scale disturbances on the quasi-stationary planetary-scale monsoon trough over the Indian region during the summer monsoon season (June to September). While some depressions form *in situ*, others can be related to weak pressure waves travelling from as far east as the south China sea¹. Regardless of their origin the depressions intensify over the Bay of Bengal and move along a well defined track from the north Bay of Bengal to western-northwestern India. The problem of the westward movement of the monsoon depressions has intrigued Indian meteorologists for a long time²⁻⁴. However, these early explanations were highly heuristic. Theoretical instability analyses⁵⁻⁹ of the basic monsoon current offer explanations regarding the initial growth of these disturbances. Many of these analyses⁵⁻⁷ show even an eastward phase speed of the gravest mode in the lower troposphere. Lindzen *et al.*⁸ examined asymptotic response to local perturbations and obtained a barotropically unstable mode in the lower troposphere with realistic scale and phase speed but having much too slow a growth rate compared to that of the monsoon depressions. With the inclusion of cumulus heating, Moorthi and Arakawa⁹ show that the gravest baroclinic mode of the monsoon current has scale, growth rate and phase speed that compare well with those of the monsoon depressions. However, the results of the linear instability analyses are relevant only for the initial growth of the disturbances. Numerical weather prediction models have been invoked in the recent years^{10,11} to simulate the westward movement of the depressions. At one end of the spectrum of such models, non-divergent barotropic models fail to produce any movement, while at the other end a multi-level primitive equation model with parameterization of cumulus convection is able to simulate reasonably realistic westward movement. As observed, the largest precipitation occurs in the western sector of the depression in such a model¹². However, the complexity of these models makes it difficult to focus on the physical process responsible for the WNW movement of the depressions. In this report, I show that the occurrence of maximum precipitation to the west could be responsible for the westward movement of the depression and suggest why the maximum precipitation occurs to the west of the disturbance.

Based on all the empirical, theoretical and modelling studies cited above, I envisage the following model regarding the initial

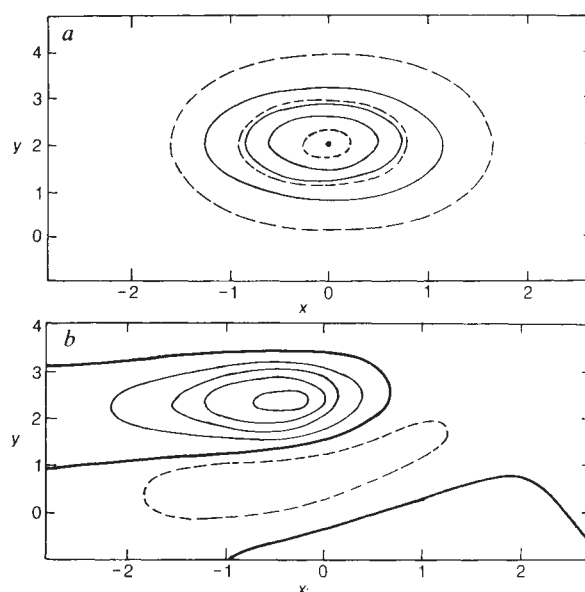


Fig. 1 Steady-state response due to a heat source centred around $x = 0$ and $y = 2$, with $\alpha^2 = 1$. *a*, The heat source (solid contour), contour interval = 0.025, no zero contour and vertical velocity (dashed contour), contour interval = 0.04. The vertical velocity is negative outside the last contour. *b*, The vorticity field, contour interval = 0.1, negative contours are dashed and the thick solid line is the zero contour.

growth and movement of the depressions. The initial growth is due to hydrodynamic instabilities. However, the growth rate can be explained only if some form of cumulus heating is included. So, I assume that the cyclonic vorticity associated with the initial instability organizes cumulus convection, and the latent heat release associated with it results in tropospheric heating. The atmosphere then responds to this heat source and the following feedback mechanism is proposed for its future developments.

Recent advances in our understanding of heat induced tropical circulation^{13,14} have suggested that the stationary atmospheric response due to an isolated heat source at such a latitude (20°N) is mainly in the form of a stationary Rossby-gravity wave to the west of the heat source¹⁴. The stationary Kelvin wave response to the east of the heat source is highly evanescent at such a latitude. This produces a strong cyclonic low-level vorticity maximum in the WNW sector of the disturbance. Therefore, frictional convergence of moisture will be highest in this sector. Also, the initial heating is associated with a mass convergence (or vertical velocity) which is collocated with the heat source. The maximum resultant moisture convergence will be at the maximum of resultant low-level upward motion (frictional + heat induced). Due to the strong cyclonic vorticity response associated with the Rossby-gravity wave to the west of the heat source this position occurs in the WNW sector. As a result, maximum moisture convergence and the heat source shifts to the WNW sector at a later time. The feedback process repeats all over again leading to the WNW movement of the monsoon depressions. I demonstrate this using a very simple dynamical model.

Our model consists of the linear shallow water equations on an equatorial beta-plane, namely

$$\begin{aligned} u_t + U_0 u_x - \beta y v + \phi_x &= -A u \\ v_t + U_0 v_x + \beta y u + \phi_y &= -A v \\ \phi_t + U_0 \phi_x + gH(u_x + v_y) &= -B\phi - Q \end{aligned} \quad (1)$$

where u , v , ϕ , Q are the zonal velocity, meridional velocity, pressure perturbation and source or sink of mass respectively. U_0 is a constant basic flow and A and B represent inverse time

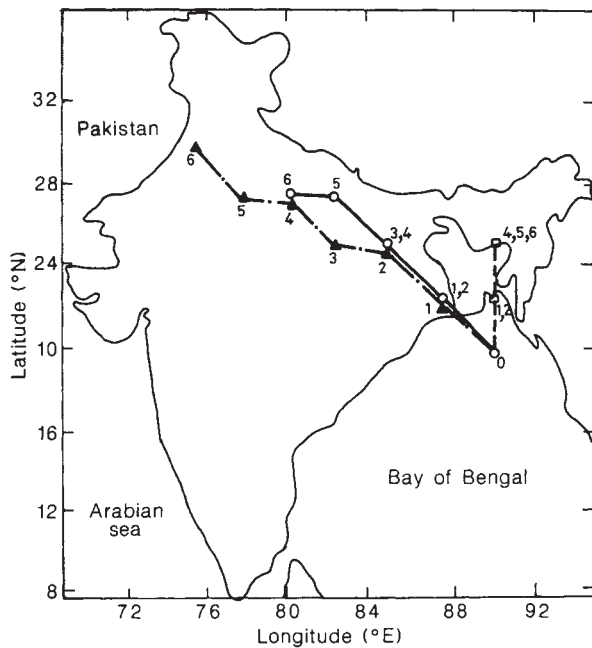


Fig. 2 Movement of the heat source and that of the depression. Originally the heat source was at the positions marked 0. Other numerals represent position of the heat source on subsequent days. Parameters are $U_0 = 0$, $A = B = Q_0 = 0.1$, $\alpha^2 = 2$ (---), 1 (—) and 0.5 (—•—).

scales for Rayleigh friction and newtonian cooling. The philosophy of using such a model to study the forced tropical circulation is discussed in detail by Gill¹³. The relevance of such a model rests on the assumption that the driving heat source for large-scale tropical motion is the net effect of the cumulus convection and that the small-scale cumulus clouds also produce enough mixing to result in an effective dissipation for the large scale flow. Chang¹⁵ found that the dissipation had the effect of inhibiting vertical propagation and eliminating waves with small vertical wavelengths. It is shown¹⁶ that in the tropics the vertical structure of the forced response closely follows the vertical structure of the forcing. Because the latent heating due to convection in the tropics has a single maximum around 500 mb¹⁷, the first baroclinic mode with an equivalent depth $H = 400$ m is a reasonable approximation to this heating function. Thus, the basic assumption in our model is that the vertical structure of the response is mainly in the first baroclinic mode with $H = 400$ m. The atmospheric long-gravity-wave speed for our problem, $C = (gH)^{1/2}$ is about 63 ms^{-1} . We nondimensionalize all the variables in equation (1) using a length scale $L = (C/2\beta)^{1/2}$ and timescale $T = (1/2C\beta)^{1/2}$. For our problem $\beta = 2.29 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$. Therefore, L is approximately 10 degrees latitude and T is about 0.22 days.

Equation (1) is solved numerically using the staggered Arakawa-C grid for horizontal differencing and a leapfrog scheme for time marching. A rectangular domain between $-10 \leq x \leq +10$ and $-2 \leq y \leq 6$ is used with $\Delta X = \Delta Y = 0.25$. The north and south boundaries are assumed to be rigid walls with no slip condition. At the east-west boundaries a periodic boundary condition is used. The parameters used are $A = B = 0.1$ and $U_0 = 0$. The heat source Q is assumed to be given by

$$Q = Q_0 \exp[-\alpha^2(x - x_0)^2 - \alpha^2(y - y_0)^2]. \quad (2)$$

so that (x_0, y_0) represents the centre of the heat source and α represents the inverse half width. Figure 1a shows such a heating distribution with $Q_0 = 0.1$, $x_0 = 0$, $y_0 = 2.0$, $\alpha^2 = 1$ and the vertical velocity (W) field of the steady response which is attained roughly after $T = 15$. In our model this vertical velocity is proportional to low-level mass convergence, $W = -(u_x + v_y)$. For the

same steady state the vorticity field (ζ_0) is shown in Fig. 1b. Note that, while the heating induces vertical velocity (or mass convergence) which is collocated with the heating field itself, the centre of the cyclonic vorticity induced by the heating is located in the WNW sector of the heat source. This cyclonic vorticity will be associated with frictional convergence of moisture proportional to $W_F = (\delta_E/2D) \zeta_0$, where δ_E is the Ekman depth and D is a vertical length scale. In the steady state the maximum moisture convergence will take place at the maximum of $(W_F + W)$. For the steady state described above assuming $\delta_E/2D = 0.25$, this position is $(-0.25, 2.25)$. I postulate that this will be the centroid of the new heat source. In response to this new heat source the centroid will again move to the WNW and the process will continue as long as there is enough moisture for maintaining the convective heat source. This, I believe, is the primary mechanism that causes the observed maximum precipitation in the western sector of the depression and its westward movement.

With the simple dynamical model, I simulate this by first allowing the atmosphere to respond to the initial heat source around 20°N that is, $x_0 = 0$, $y_0 = 2.0$ up to $T = 15$ and then updating the heat source using equation (2) every time step with $(x_0, y_0) = \text{maximum of } (W_F + W)$. The position of the heat source and hence that of the depression on subsequent days are shown in Fig. 2 for $U_0 = 0$. Although the WNW movement of the depressions is a rather robust feature of our model, the actual track and the speed of movement depend to some extent on the choice of parameters α , $\delta_E/2D$ and so on. Figure 2 shows that larger depressions (half width $\geq$ equatorial Rossby radius of deformation) move easily to the WNW direction while smaller depressions may move to the north. This is a qualitative prediction of my simple dynamical model that may be tested when data on the sizes of the depressions are available. We also find that for a given size of the depression, $\delta_E/2D$ must be greater than a certain threshold for the WNW movement of the depressions. For $A = B = Q_0 = 0.1$ and $\alpha^2 = 0.5$, the threshold is $\delta_E/2D = 0.15$. This means that the boundary layer must be deeper than a critical value for the westward movement of the depressions. If we include an easterly mean flow, the depressions move faster in a more westward direction while if we include a westerly mean flow, the depression would move more towards the north and north-east. In reality, however, the depressions occur on the quasi-stationary planetary-scale monsoon trough. On the trough $U_0 \approx 0$. The basic flow changes from weak westerly, south of the trough, to weak easterly in the north of the trough. However, to study the movement of the heat source in the presence of a spatially varying basic flow is outside the scope of the model. Moreover, I want to emphasize that the aim of this work has been merely to demonstrate the feasibility of the proposed physical mechanism and not to simulate any given track of a depression.

The mechanism proposed in this report predicts maximum precipitation in the northwestern sector of the disturbance while some observations¹⁸ indicate that it occurs in the southwestern sector. This may be related to our assumption that in my model the moisture supply to the heat source is determined mainly by the frictional convergence at the lowest layer. In reality, however, the horizontal advection of moisture and other dynamical processes may be important. These processes combined with the frictional convergence in the northwestern sector seem to produce the maximum convergence of moisture in the southwestern sector. This would also explain the more westward movement of the monsoon depressions than that predicted by the model. Secondly, in my simple model all the complex dissipative processes are parameterized by linear Rayleigh friction (A) and newtonian cooling laws (B). The values of A and B used in our model represent dissipative process with a timescale of 2–3 days. This timescale is rather close to the depression timescale itself. The qualitative result of our analysis remains unchanged even if we increase A and B to represent dissipative processes having

shorter timescales. Such an *ad hoc* increase in *A* and *B* may even be justified as certain diagnostic studies¹⁹ point out that increased vertical mixing by the cumulus clouds is important in the monsoon depressions.

I thank Professor M. Sankar Rao for his help in designing the numerical model.

Received 24 June; accepted 30 December 1986.

1. Saha, K., Sanders, F. & Shukla, J. *Mon. Weath. Rev.* **109**, 330-345 (1981).
2. Koteswaram, P. & George, C. A. *Indian J. Met. Geophys.* **9**, 9-22 (1958).
3. Koteswaram, P. & Bhaskar Rao, N. S. *Aust. Met. Mag.* **41**, 62-75 (1963).
4. Rao, K. V. & Rajamani, S. *Indian J. Met. Geophys.* **21**, 187-194 (1970).

5. Shukla, J. *J. atmos. Sci.* **35**, 495-508 (1978).
6. Goswami, B. N., Keshavamurty, R. N. & Satyam, V. *Proc. Indian Acad. Sci. (Earth. planet. Sci.)* **89**, 79-97 (1980).
7. Mishra, S. K. & Salvekar, P. S. *J. atmos. Sci.* **37**, 383-394 (1980).
8. Lindzen, R. S., Farrell, B. & Rosenthal, A. J. *J. atmos. Sci.* **40**, 1178-1184 (1983).
9. Moorthi, S. & Arakawa, A. *J. atmos. Sci.* **42**, 2007-2031 (1985).
10. Krishnamurti, T. N. *et al. J. met. Soc. Japan* **54**, 208-224 (1976).
11. Krishnamurti, T. N. *Adv. Geophys.* **28B**, 283-333 (1985).
12. Krishnamurti, T. N., Molinari, J., Pang, H. L. & Wong, V. *Pure appl. Geophys.* **115**, 1357-1372 (1977).
13. Gill, A. E. *Q. Jl R. met. Soc.* **106**, 447-462 (1980).
14. Heckley, W. A. & Gill, A. E. *Q. Jl R. Met. Soc.* **110**, 203-217 (1984).
15. Chang, C. P. *J. atmos. Sci.* **34**, 901-910 (1977).
16. Geisler, J. E. & Stevens, D. E. *Q. Jl R. Met. Soc.* **108**, 87-93 (1982).
17. Yanai, M., Esbensen, S. & Chu, J. H. *J. atmos. Sci.* **30**, 611-627 (1973).
18. Godbole, R. V. *Tellus* **29**, 25-40 (1977).
19. Daggupaty, S. M. & Sikka, D. R. *J. atmos. Sci.* **33**, 773-792 (1977).

Evidence for transform margin evolution from the Ivory Coast-Ghana continental margin

Jean Mascle & Emmanuel Blarez

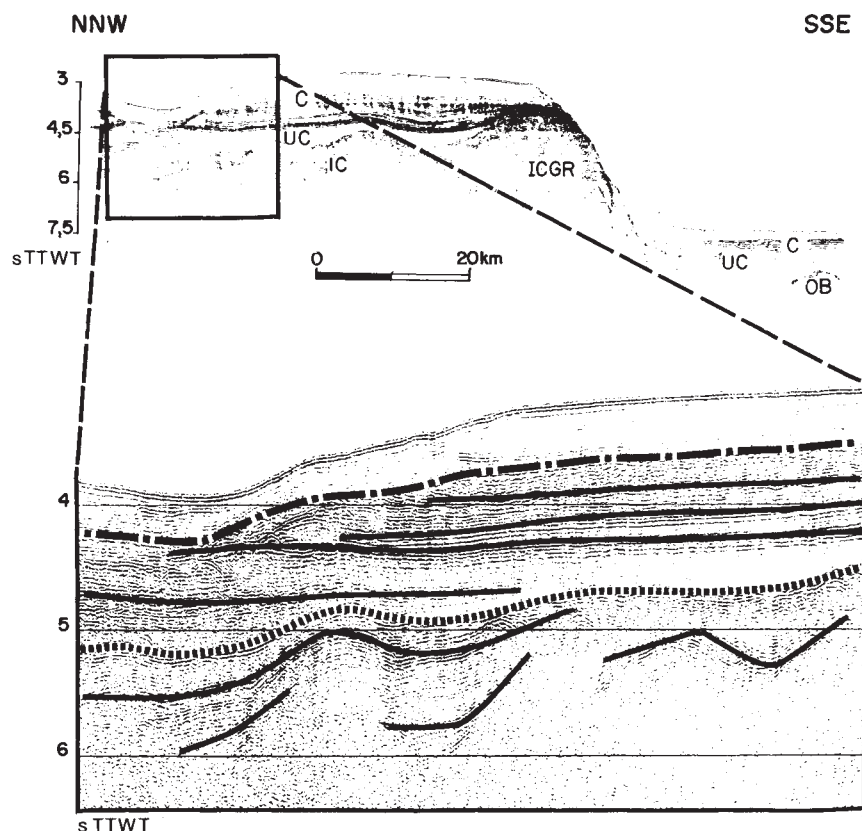
Laboratoire de Géodynamique sous-marine, CEROV, BP 48-06230 Villefranche-sur-mer, France

The relative motions between lithospheric plates can be reduced to three major types: convergence, divergence and transform motion. Convergence leads to an active margin (and eventual collision); divergence results in the construction of a rifted (passive) margin; transform motion generates a transform (or sheared) margin. Although passive margins have been extensively studied^{1,2}, and many models have been proposed for their origin and subsequent evolution³⁻⁵, little is known about transform margins, with the exception of a few studies of their morphologies^{6,7}, shallow

structures or crustal sections^{8,9}. Here we present the main conclusions derived from a recent study of the northern Gulf of Guinea margins, particularly off the eastern Ivory Coast and Ghana, where the continental margin is one of the best-preserved examples of an extinct transform margin¹⁰⁻¹². Our observations support a four-stage model for transform margin evolution. Tectonically active transform contacts, first between normal continental crusts and then between thinned margins, induce characteristic structures such as pull-apart grabens and shear folds. The next stage, in which thermal exchange between oceanic and continental lithospheres controls a complex subsidence, is followed by the transition to a true intra-oceanic fracture zone.

The interpretations presented here are based mainly on data collected during a survey run in 1983 off the Ivory Coast and Ghana¹³. About 3,300 km of continuous geophysical profiles (3.5 kHz, magnetism and single channel seismic reflection) were obtained along the Ghanaian continental slope and over the eastern Ivory Coast margin. Ten dredge stations were occupied chiefly along the Ghanaian slope. Detailed analysis of the results have been published or are in the press^{13,14}. Our data have been

Fig. 1 Example of a single channel seismic line across the western Ghanaian continental margin from the middle slope (on the north-north-west (NNW)) to the Guinean abyssal plain (south-south-east (SSE)) (see location on Fig. 2). Vertical scale in seconds two-way travel time (s TTWT). In this area a deeply sedimented basin represents the prominent feature of the margin. Towards the south the basin appears dammed against an acoustic basement ridge, the so-called Ivory Coast-Ghana Ridge (ICGR) facing, through a steep slope, the deep Gulf of Guinea abyssal plain. We note, within the basin sedimentary cover, the presence of two distinct seismic sequences separated by a strong unconformity. The upper sequence consists of a series of well-layered and almost flat-lying seismic reflectors thought to be of late Cretaceous to Pleistocene age (an erosional unconformity of probable Oligocene age is well expressed in this sequence). The lower sequence is made of obviously deformed reflectors, particularly well seen in the enlargement. A series of anticline and syncline features have developed within this sequence attributed to early lower Cretaceous deposits. C, Cenozoic, UC, Upper Cretaceous, LC, Lower Cretaceous and OB, oceanic basement.



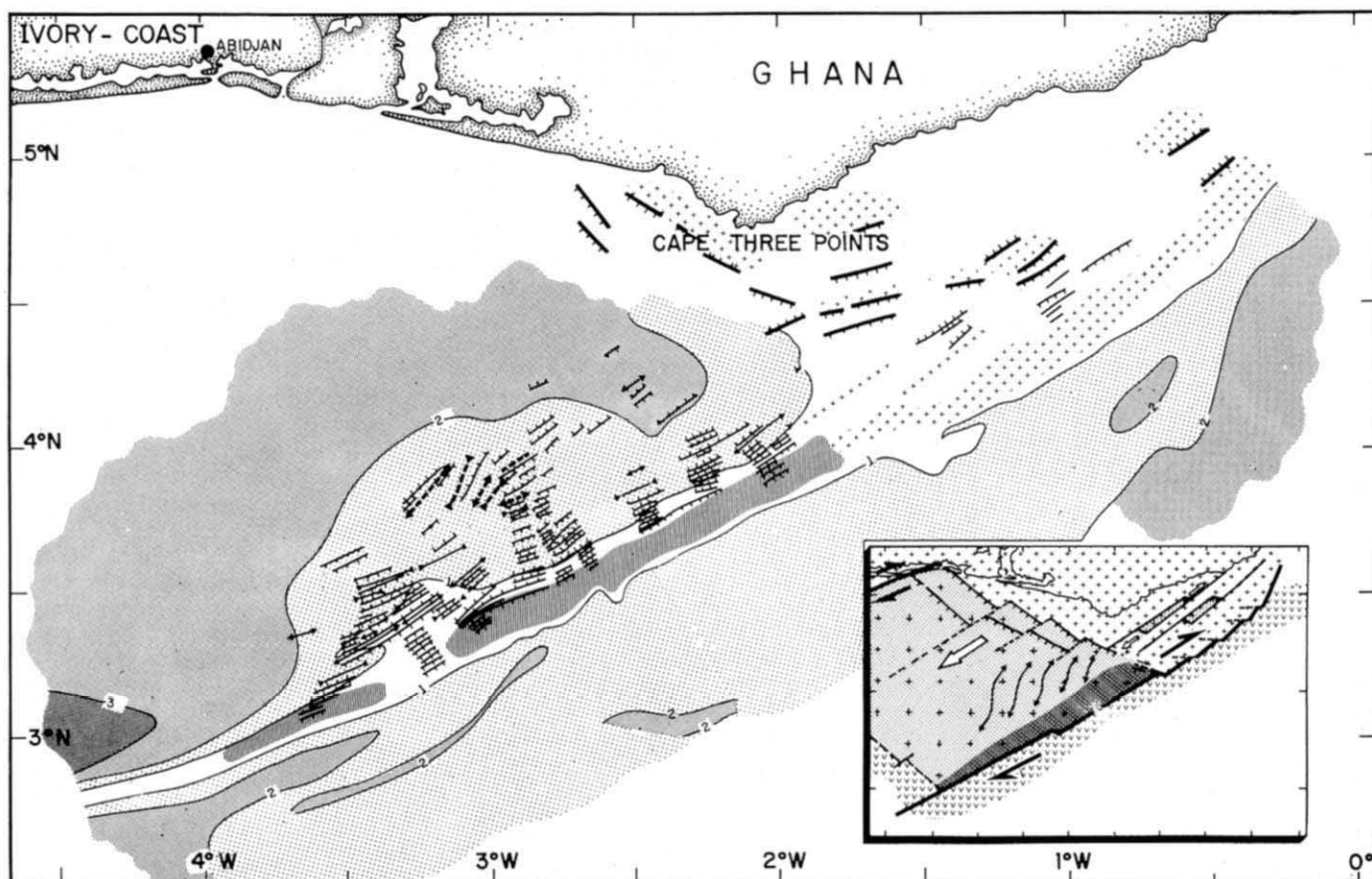


Fig. 2 Schematic structural sketch of the Ivory Coast-Ghana margin (insert shows the structural interpretation). We have mapped the main tectonic features observed at the level of the lower seismic sequence. The Ghanaian platform comprises a series of structural highs (+++) and asymmetric grabens disconnected by fault lineaments (---), the ICGR (stippled) corresponds to an acoustic basement rise and may be partly of sedimentary origin; in the deep Ivory Coast Basin (ICB), we note the presence along the ICGR of a belt of faulted and folded sediments. Light, medium and dark dotted areas indicate respectively sedimentary thickness between 1 and 2, 2 and 3, and more than 3 s in two-way travel time (s TWTT). In the insert, the structural interpretation of the margin shows the deep ICB on a thinned continental crust characterized by shear sedimentary folds developing as a consequence of a right lateral transform motion. The Ghanaian slope is interpreted as an 'en echelon' structure while the Ghanaian platform grabens correspond to probable pull-apart basins emplaced between secondary strike-slip fault lineaments.

supplemented by a few unpublished multi-channel seismic (MCS) lines which have allowed us to attempt seismic stratigraphic correlations with boreholes located on the Ivory Coast and Ghana shelves.

Figure 1 shows an example of a single-channel seismic line recorded across the margin from the middle slope to the Guinea abyssal plain. The continental margin appears mainly characterized by a thick sedimentary slope basin dammed southward against a basement-like ridge, the Ivory Coast-Ghana Ridge (ICGR) facing the oceanic abyssal plain. The marginal ICGR has developed as a 30 km wide and quasilinear (more than 200 km long) east-north-east to west-south-west oriented feature in continuity with the Ghanaian continental slope. The sedimentary infilling of the slope basin, the deep Ivory Coast basin (ICB), contains two major lithoacoustic units (Fig. 1). The upper one consists of a series of well-layered reflectors and the lower unit is made of obviously deformed seismic reflectors. The seismic units are separated by a strong unconformity detected over the entire margin as well as on the Ghanaian platform.

A closer inspection indicates that the upper unit contains several unconformities including a major one of probable Oligocene age¹⁵. Correlations through MCS lines with offshore boreholes indicate a late Cretaceous to Cenozoic age for this upper seismic unit. The lower unit, where synclines-anticlines type structures are locally detected, is also by correlation, attributed to early Cretaceous sediments. Towards the ICGR,

this deep unit progressively pinches out, but it is impossible to know if this is due only to a progressive sedimentary thinning or to an increasing deformation leading to an acoustic basement like signature. An early Cretaceous age for this unit is coherent with offshore Ivory Coast boreholes indicating the presence within the drilled sections of a major upper Albian to lower Cenomanian unconformity.

If these attributions are correct, the margin sedimentary cover can be divided into an early Cretaceous to upper Albian tectonically deformed sequence and a Cenomanian to Pleistocene unconformable cover.

In Fig. 2 we have constructed a structural sketch of the ICGR margin showing the main tectonic features detected within the lower unit. On the Ghanaian shelf this mapping remains hypothetical because of both the shallow depth and the lack of available MCS lines. We observe evidences of north-north-west dipping reflectors within asymmetric basins disconnected by apparently north-north-east to south-south-west trending fault systems. Boreholes on the Ghanaian platform indicate the presence of thick lower Cretaceous sediments resting directly on Precambrian or Palaeozoic units. The dredge hauls attempted along the Ghanaian slope, have yielded only rock fragments which can be correlated with African basement units¹⁴. Towards the south-west, the Ghanaian platform and the Ghanaian slope propagate respectively through the ICB and ICGR. As previously mentioned the ICGR does not show, on seismic grounds,

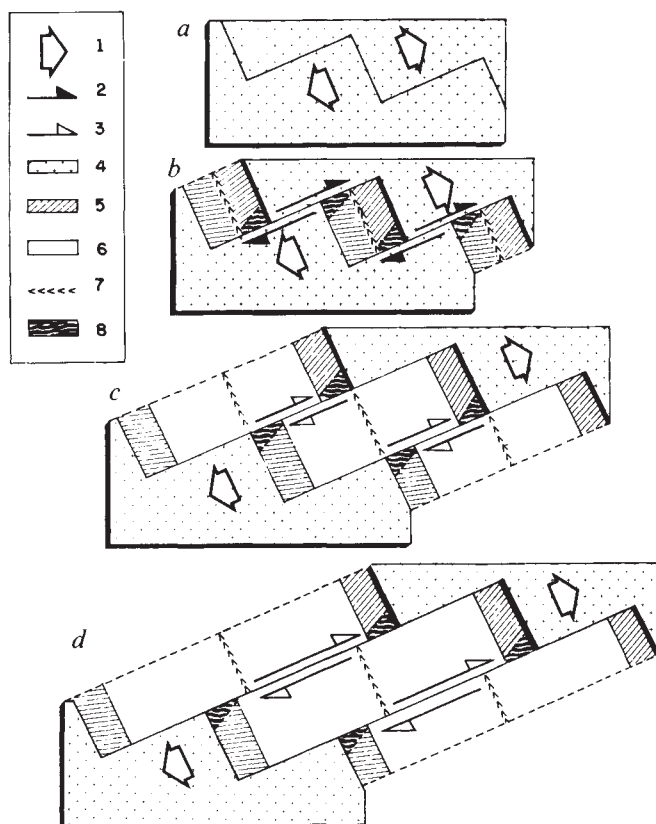


Fig. 3 A simplified model for the evolution of transform margins based on the example of Ivory Coast and Ghana margins (see text for detailed explanation). *a*, Continent to continent active transform contact. *b*, Continent to continental margin (thinned crusts and sedimentary cover) active contact, creation of lateral marginal ridge and associated tectonics by shearing. *c*, Progressive drift of the margin along a hot accretionary centre. Thermal exchange between a continental lithosphere and an oceanic lithosphere leading to vertical readjustments. *d*, Mature stage. Thermal subsidence evolution. 1, Divergence, 2, transform motion between continental crusts; 3, transform motion between oceanic crusts (fracture zones); 4, thick continental crust; 5, thinned continental crust; 6, oceanic crust; 7, mid-oceanic ridge axis; 8, marginal ridge and related tectonics (thinned crustal blocks and deformed sedimentary wedge).

any significant reflector. Dredge stations made along its southern slope have only produced barren sedimentary rocks (mostly siltstones, sometimes folded) indicating that the feature includes 'deformed' sediments. North of the Ridge, within the deep ICB, the different tectonics features (faults and folds) constitute a belt underneath and parallel to the Ridge axis east-north-east to west-south-west about 60 km wide, where the anticline-syncline systems show an apparent north-north-east to south-south-west axial direction.

The data briefly presented above indicate: that apparent compressive deformations occur within the sedimentary cover of parts of the Gulf of Guinea margin; that these deformations characterize strata, attributed with a good degree of confidence, to early Cretaceous times and finally that a tectonic unconformity separates the lower seismic sequences from upper units attributed to late Cretaceous-Cenozoic deposits.

In the past years models of the evolution of the South Atlantic, based on magnetic lineations¹⁶ fracture zone trends¹⁷ or combinations of both¹⁸ have been proposed. All imply an ocean opening starting around 130–125 Myr BP and an oblique (or transform) motion within the equatorial area, particularly along the northern Gulf of Guinea. Taking into account both the geometry of opening and average spreading rates of about 2 cm

yr⁻¹, we deduce that oceanic accretion was not active within the future Gulf of Guinea before Alptian-Albian time. As a direct consequence it can be argued that a continent-to-continent contact (or a continental margin contact) was effective between Brazil and Ivory Coast-Ghana until late Albian-early Cenomanian. It is likely then that the structures of both the Ghanaian platform and slope originated during early Cretaceous together with the progressive creation of a short segment of rifted margin (the present-day south-eastern Ivory Coast margin) and that the tectonic deformations detected within the ICB and along the ICGR originated in relationship with the progressive sliding of two continental masses (and adjacent margins) one against the other. The late Albian-early Cenomanian unconformity may then indicate the timing of the final separation between the Brazilian and African continental crusts. In other words, the tectonic activity observed along this margin results essentially from an active transform contact between two continental crusts and would have ceased when the junction between the oceanic crust on both sides of the main shear zone was effective.

Off the Ivory Coast, because of the length of the initial offset, one should expect, until late Cretaceous times, an active tectonic contact between the margin (and its deformed sedimentary cover) and a hot centre (an accretionary centre) drifting progressively along the shear zone. Such contact should have induced strong thermal contrasts and subsequent crustal readjustments. A few sedimentary onlaps within the late Cretaceous sedimentary cover of the ICGR and the relative high elevation of the Ridge may be direct consequences of this phenomenon.

According to the rate of oceanic opening, the continental margin (including the ICGR) should have passed laterally along a southern spreading centre around Santonian times. Since that time, the Ivory Coast-Ghana margin has probably evolved as other more classic passive margins, even if thermal effects, due to contrasts between two different (oceanic to the south, continental to the north) lithospheres can be suspected.

We have indicated above that most of the morphostructural characteristics of the Ivory Coast-Ghana margin (including the creation of a marginal ridge), have probably been obtained during early transcurrent evolution. Since late Albian-Cenomanian the margin has not been submitted to any major geodynamic rearrangement. In other words, the Ivory Coast-Ghana margin can be regarded as a typical fossil transform margin. If this is correct, we propose as a direct consequence of this example, that transform margins can be characterized by four main stages of evolution (see Fig. 3): (1), a continent to continent active contact, (2), a continent (or continental margin) to continental margin active contact, (3), an oceanic crust to continental margin active contact, (4), finally a mature stage during which the transform contact is active only between two areas of oceanic crust (an oceanic fracture zone is created).

During the first stage (contact between two continental masses) two, probably thick, crusts are in a transform contact; their relative motion should generate apparent brittle deformations within the upper crust and probable ductile deformations at depth⁸. Depending on the prevailing constraints, pull-apart basins, structurally rotated blocks, are created; the present California day¹⁹ or the northern Red Sea^{11–20,21} are potential examples of such a contact.

During the second stage, small segments of divergent margins are progressively created on both sides of the main shear zone. Active transform contacts occur between a stretching continental crust and a thicker continental crust. Within the newly created rifted basin (surrounded by land masses and tectonically active fault zones), both the sedimentation and the subsidence are probably particularly high. The basin sedimentary cover is progressively deformed, mainly along the area locked between the two sliding crusts; faults, shear folds, marginal ridge (possibly including a highly deformed wedge of sediments) are emplaced.

We postulate that the end of the active contact between the two continental crusts leads to both the end of synsedimentary deformations and to a posterior tectonic unconformity (the late Albian-early Cenomanian unconformity off Ivory Coast). This unconformity can be regarded as an equivalent of the post break-up unconformity of classic rifted margins.

The newly created margin should then progressively drift along a hot crust (the accretionary centre). This evolving contact, between a hot centre and a continental lithosphere, should necessarily induce by thermal exchanges some kinds of crustal readjustments. We believe that this can be recorded in a succession of sedimentary onlaps and in a marked slow down (or even an end) of normal thermal subsidence process. Actually the relative high elevation of marginal ridges can be a direct consequence of this stage.

After this period (progressive shift of the margin along an accretionary centre), the transform continental margin is directly adjacent to a hot (but cooling) oceanic lithosphere; its subsidence behaviour should then be comparable to that of classic rifted margins.

We thank our colleagues of the GEMC for discussion. The data were collected in 1983 during a cruise on board the RV *Le Suroit* (IFREMER). Financial support was given by CNRS (ATPGGO).

Received 20 October 1986; accepted 19 January 1987.

1. Burk, C. A. & Drake, C. L. *The Geology of Continental Margins* (Springer, Berlin, 1974).
2. Watkins, J. D. & Drake, C. L. (eds) *Mem. Am. Ass. Petrol. Geol.* 34 (1982).
3. McKenzie, D. P. *Earth planet. Sci. Lett.* 40, 25-32 (1978).
4. Le Pichon, X. & Sibuet, J. C. *J. geophys. Res.* 86, 3708-3720 (1981).
5. Wernicke, B. *Can. J. earth. Sci.* 22, 108-125 (1983).
6. Mascle, J. *Anais Acad. bras. Cienc.* 48, 191-197 (1976).
7. Lonsdale, P. *Bull. Am. Ass. Petrol. Geol.* 69, 1160-1180 (1985).
8. Scrutton, R. A. *Tectonophysics* 59, 293-305 (1979).
9. Scrutton, R. A. (ed.) *Dynamics of Passive Margins* 133-140 (Am. Geophys. Un., Washington, DC, 1982).
10. Mascle, J., Mougnot, D., Blarez, E., Marinho, M. & Virlogeux, P. *J. Geol.* (in the press).
11. Blarez, E. *thesis Univ. Paris VI* (1986).
12. Mougnot, D., Virlogeux, P., Vanney, J. R. & Malod, D. *Bull. Soc. géol. Fr.* 8, 419-422 (1986).
13. Blarez, E. & Mascle, J. *Les Marges Transformantes Ivoir-Ghanéenne et Sud Guinéenne. Campagnes Océanographiques Françaises No. 3* (IFREMER, Brest, 1986).
14. Blarez, E. *et al. Bull. Soc. géol. Fr.* 8, 419-422 (1986).
15. Simon, P. & Brou, A. *Bull. Soc. géol. Fr.* 26, 1117-1125 (1984).
16. Rabinowitz, P. D. & Labrecque, J. *J. geophys. Res.* 84, 5973-6002 (1979).
17. Le Pichon, X. & Hayes, D. E. *J. geophys. Res.* 76, 6283-6293 (1971).
18. Sibuet, J. C. & Mascle, J. *J. geophys. Res.* 83, 3401-3421 (1978).
19. Crowell, J. C. *Spec. Publ. Soc. econ. Paleont. Miner.* 19, 293-303 (1974).
20. Ben Avraham, Z. *J. geophys. Res.* 90, 90-131, 703-726 (1985).
21. Ben Avraham, Z., Almagor, G. & Garfunkel, Z. *Sedim. Geol.* 23, 239-267 (1979).

Lead isotope evidence for the structure of the Rockall dipping-reflector passive margin

Andrew C. Morton* & Paul N. Taylor†

* British Geological Survey, Keyworth, Nottingham NG12 5GG, UK

† Department of Earth Sciences, University of Oxford, Oxford OX1 3PR, UK

The passive continental margin south-west of Rockall Plateau is characterized by a thick sequence of oceanward-dipping seismic reflectors. During Leg 81 of the Deep Sea Drilling Project, these reflectors were sampled at Site 553 and proved to consist almost exclusively of basalt. Here we present lead isotope data which indicate that these basalts may have been contaminated by ancient uranium-depleted continental crust, or alternatively, derived from a sub-continental lithospheric mantle source. In either case, the implications are that the basalts of the south-west Rockall Plateau formed by eruption through and onto continental basement, not by 'subaerial seafloor spreading'. This conclusion is in accord with gravity models of the area, which predict stretched continental crust beneath the dipping reflector sequence.

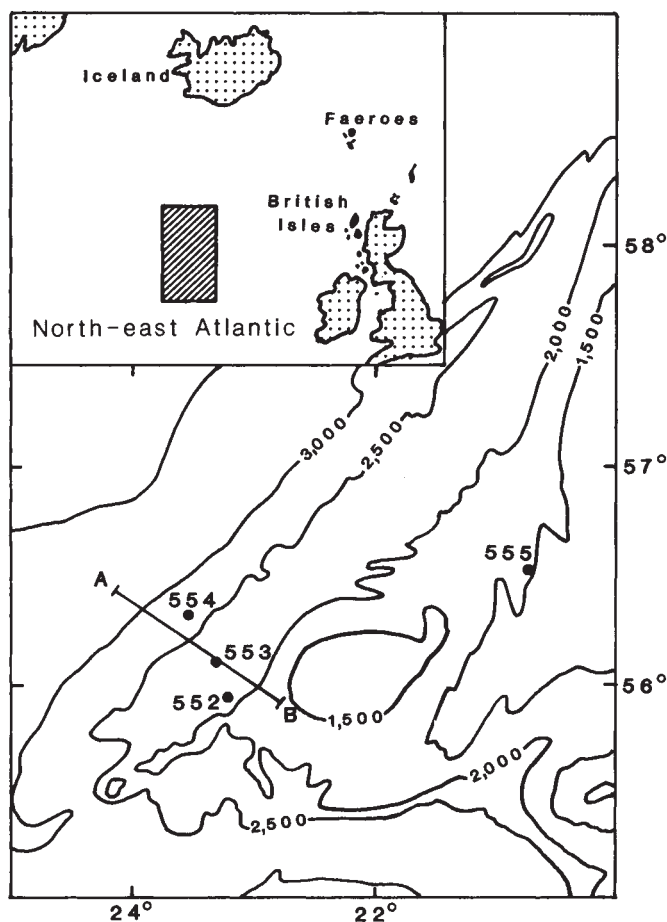


Fig. 1 Location of DSDP Site 553 in the North-east Atlantic and the location of the seismic reflection profile shown in Fig. 2.

Seismic reflection profiles across the passive continental margin south-west of Rockall Plateau, North-east Atlantic (Fig. 1) have revealed the existence of a group of oceanward-dipping reflectors (Fig. 2), a structure that differs considerably from that found on block-faulted Biscay-type margins¹. Similar structures have been observed elsewhere in the North-east Atlantic, off East Greenland^{2,3}, around the northern and western margins of the Faeroes Plateau⁴, and in the Voring Plateau area^{5,6}. Hinz⁵ has recognized similar features on other passive continental margins (off south-west Africa, north-west Australia, and the east coast of the USA, and in the Weddell Sea, off Antarctica). These so-called dipping-reflector margins appear to be at least as widespread as Biscay-type margins, and may prove to be the predominant type⁷, but their origins are less well understood.

Five models have been put forward to explain the formation of dipping-reflector sequences (DRSs).

Model 1. Results of drilling in the south-west Rockall Plateau area during DSDP Leg 48 led Roberts *et al.*¹ to suggest that the DRSs were generated by progradation of shallow-water and deltaic clastic sediments, interbedded with extrusives, during rifting. Mutter *et al.*⁶ comment that, in several respects, the structure mimics the prograding clinoforms of a deltaic complex. However, DSDP Leg 38 proved the top of the DRS in the Voring Plateau area to be basaltic⁸, and DSDP Leg 81 also proved this to be the case in the south-west Rockall Plateau area.

Model 2. Hinz⁵ proposed that the DRS was generated by eruption of basalts and pyroclastics during rifting from points along the line of eventual continental separation. In model 2, therefore, the DRS lies above thinned continental crust, with the continent-ocean boundary (COB) located at the landward limit of normal oceanic crust.

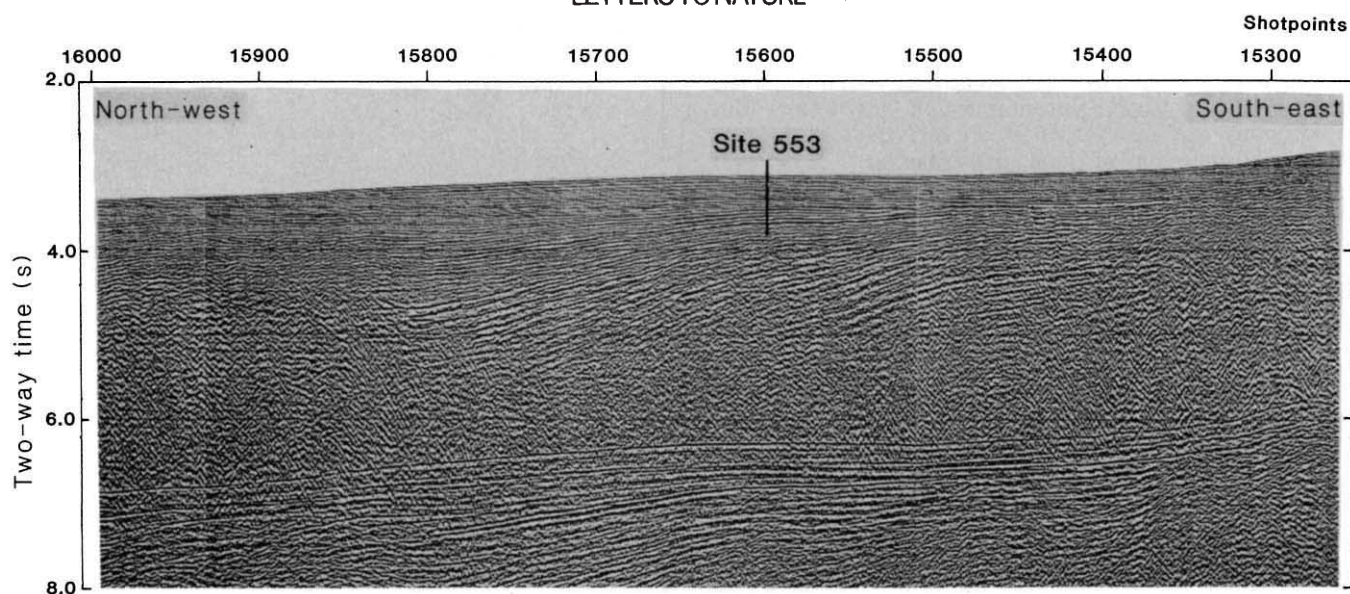


Fig. 2 Seismic reflection profile across the passive margin south-west of the Rockall Plateau, showing the wedge of oceanward-dipping reflectors and location of Site 553. From ref. 9.

Model 3. Proposed by Roberts *et al.*⁷, this is a variant of model 2. In this model, the DRS is again underlain by thinned continental crust, but with dyke intrusion throughout the zone of stretching. Gravity modelling suggests that continental crust is present under the DRS in the south-west Rockall Plateau¹⁰; however, gravity models are not definitive, as solutions are not unique. **Model 4.** From their work on the Voring Plateau, Mutter *et al.*⁶ proposed that the DRS was generated by 'subaerial seafloor spreading'. They proposed that the Voring Plateau Escarpment, located landward of the DRS, marks the COB. There are a number of problems associated with such a position. Firstly, Smythe *et al.*¹¹ suggested that such a position is inconsistent with gravity data, although this has been disputed by Talwani *et al.*¹². Secondly, the Faeroe-Shetland Escarpment, the southerly counterpart of the Voring Plateau Escarpment, runs to the east of the Faeroe Islands^{4,11}, and lead isotope work on the Faeroes basalt sequence has proved beyond reasonable doubt the existence of ancient uranium-depleted continental crust under the region¹³. Thirdly, the Voring Plateau Escarpment has no counterpart in the south-west Rockall Plateau area. Fourthly, drilling at ODP Leg 104 Site 642 on the Voring Plateau penetrated extrusives with undoubted continental crustal contamination below a uniformly basaltic dipping reflector sequence¹⁴.

Model 5. Proposed by Smythe⁴, this is a variant of model 4. This model accepts that the DRS was formed by subaerial seafloor spreading, but proposes that the earliest basalts flowed onto continental crust. The COB in model 5 is placed at the lower limit of the first reflector, where it is inferred to pass into a dyke of oceanic layer 2. This model avoids the problems associated with model 4 and appears to be consistent with ODP Leg 104 drilling results, Site 642 being located at the landward limit of the DRS¹⁴.

DSDP Leg 81 was dedicated to a transect across the DRS in the south-west Rockall Plateau area⁹. Four sites were drilled, 552–555. Site 555 was located landward of the DRS, on the Rockall Plateau itself. Hole 554 drilled the 'outer high', a structure located at the boundary between 'normal' oceanic crust and the DRS. Holes 552 and 553 were located over the wedge of dipping reflectors, and both penetrated basalt. However, Site 552 is close to the landward limit of the DRS (Fig. 2), in a position where contamination by continental lithospheric material could reasonably be expected in any of the proposed

models of DRS formation. Furthermore, there is some doubt about the stratigraphic relationships of the basalts in Hole 552: they appear to be stratigraphically higher than the DRS, and may belong to a younger flow unit. In contrast, Site 553 is situated midway across the wedge, some considerable distance from Rockall Bank, in an ideal position to test the various hypotheses by geochemical means. Hole 553A recovered 183 m of plagioclase-phyric tholeiites of oceanic affinity^{15–17}, forming the upper part of the DRS at this site.

Before this study, only a limited amount of isotope data had been made available from the basalts of Hole 553A. Macintyre and Hamilton¹⁸ reported ⁸⁷Sr/⁸⁶Sr and ¹⁴³Nd/¹⁴⁴Nd values from two samples, and Harrison and Merriman¹⁵ provided ⁸⁷Sr/⁸⁶Sr for an additional one. The meagre data give no suggestion of contamination by continental crust, although this in itself is not proof for the absence of continental crust¹⁸. In view of the successful application of lead isotope studies in the detection of crustal contamination in the pencontemporaneous Faeroes and Hebridean basalts^{12,19}, we examined the lead isotopic compositions of a suite of basalt samples from the dipping reflector interval in Hole 553A. Samples were prepared by loading the rock sample in 1 M HBr solution on to anion-exchange resin, eluting Pb with HCl. They were then loaded on single Re filaments with H₃PO₄ and silica gel activator, and were analysed on a VG Isomass 54E fully automated thermal-ionization mass spectrometer. Table 1 shows analyses, corrected for the effects of mass fractionation. Lead isotopic compositions of the basalts from Hole 553A are shown in the ²⁰⁷Pb/²⁰⁴Pb–²⁰⁶Pb/²⁰⁴Pb diagram (Fig. 3), together with the compositional field of North

Table 1 Lead isotopic compositions of basalts, DSDP Leg 81, Site 553

Core	Section	Depth (cm)	²⁰⁶ Pb/ ²⁰⁴ Pb	²⁰⁷ Pb/ ²⁰⁴ Pb	²⁰⁸ Pb/ ²⁰⁴ Pb
38	1	39–41	18.035	15.567	37.684
40	2	79–81	18.111	15.538	37.833
45	3	79–81	17.675	15.503	37.349
47	3	30–32	17.968	15.493	37.589
49	3	75–77	18.180	15.553	37.914
51	2	80–82	17.910	15.518	37.584
53	3	116–118	18.092	15.511	37.705
54	4	60–62	17.190	16.527	37.002
58	1	79–81	17.799	15.534	37.526

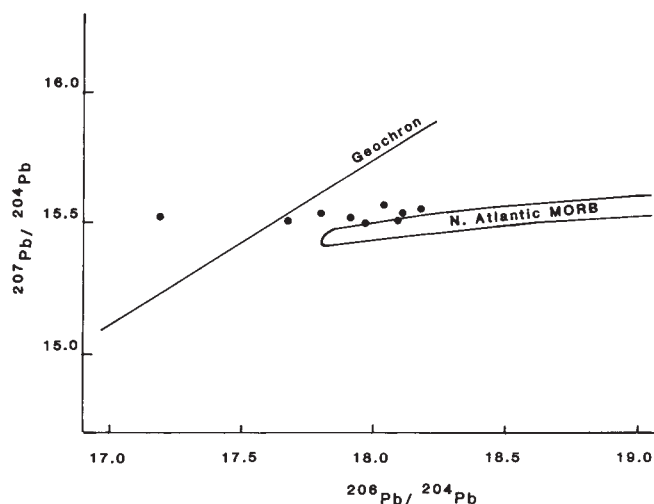


Fig. 3 $^{207}\text{Pb}/^{204}\text{Pb}$ - $^{206}\text{Pb}/^{204}\text{Pb}$ plot of basalts from DSDP Site 553, south-west Rockall Plateau.

Atlantic mid-ocean-ridge basalts (NAMORB)^{20,21} and the geochron. The data define a nearly horizontal linear array in the Pb-Pb diagram, extending from the less radiogenic end of the NAMORB field to a single relatively unradiogenic composition (553A-54-4, 60-62 cm), significantly to the left of the geochron. This latter composition would be very unusual in a basalt of purely oceanic affinities. The retardation of lead isotope evolution required to produce such a composition is characteristic of ancient continental lithospheric material, either uranium-depleted lower crust or sub-continental lithospheric mantle. Consequently we interpret the nearly horizontal array of Site 553 basalt data as the result of mixing of magmas that were derived from MORB-source mantle, either with contaminants derived from an ancient uranium-depleted crustal source, or with magmas derived from long-lived continental lithospheric mantle sources.

Alternatively, it might be supposed that the observed trend is the result of contamination with laterally derived sediment. However, as continentally derived sediments principally sample the upper, uranium-enriched continental crust, they are characterized by rather radiogenic lead: contamination of basalts with such material generates steeply inclined linear arrays in the $^{207}\text{Pb}/^{204}\text{Pb}$ - $^{206}\text{Pb}/^{204}\text{Pb}$ diagram, directed towards lead isotopic compositions more radiogenic than MORB, unlike the trend observed in Site 553 basalts.

Similarly, hydrothermal alteration is unlikely to have generated the observed trend, although all the basalts from Site 553 show some degree of alteration, mostly to clay minerals¹⁵. The effects are strongest in vesicular portions of flows and in interstitial glass, but the primary minerals are little affected. There is some evidence that the earliest stages of alteration may have been subaerial, but the main alteration is due to prolonged and pervasive contact with sea water, at temperatures probably $<150^\circ\text{C}$ ¹⁵. However, contamination with seawater lead could not account for the observed trend towards an unradiogenic lead isotopic composition, because sea water has a very low concentration of dissolved lead, about 2 pg g^{-1} (ref. 22), and furthermore its isotopic composition is more radiogenic than the Site 553 basalts.

The nearly horizontal slope of the Site 553 basalt data array would in itself be an unusual feature of an oceanic basalt series. Further, this trend does not match closely the contamination arrays of the Tertiary volcanics on Skye (contaminated with Archaean Lewisian lead¹⁹) or in the Faeroes¹²: both have steeper slopes. A somewhat complex model for lead isotope evolution

is required to generate $^{207}\text{Pb}/^{204}\text{Pb}$ ratios comparable with modern MORB-source mantle in association with markedly lower $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. This requires a source with a long history of separation from convecting mantle material, with an early stage of U/Pb enrichment relative to the main mantle reservoir, followed by a later stage with markedly lower U/Pb ratios. A possible model could be late-Archaean crust formation followed by mid-Proterozoic uranium depletion during high-grade metamorphism. Although this is not a unique solution, it is worth noting that both Laxfordian (mid Proterozoic) and Grenvillian (late Proterozoic) granulites have been recovered from the Rockall Bank²³: the lead isotope characteristics of the DRS basalts south-west of Rockall Plateau could be explained by subjacent continental lithosphere with a similar long and complex geological history.

Even without the precise details of the age characteristics of the basement to the DRS, there can be little doubt that the basalts have a component of Pb derived from subjacent ancient continental lithosphere. The presence of such material beneath the DRS is inconsistent with models 4 (Mutter *et al.*⁶) and 5 (Smythe⁴) discussed above. In both models the basalts at Site 553 would have been erupted seaward of their present location, at the point where the basalt flows are inferred to pass into feeder dykes of oceanic layer 2. However, from the lead isotope data it is apparent that continental lithospheric material is present beneath the site of eruption.

The continent-ocean boundary cannot therefore be in the position implicit in either models 4 or 5. In contrast, models 2 (Hinz²) and 3 (Roberts *et al.*⁷), which both propose that the DRSs are underlain by continental crust, are consistent with both the isotope data reported here and the gravity data¹⁰. Of the two models, model 3, which proposes dyke injection throughout the zone of extension, is favoured as being the most geologically reasonable. The position within the dipping-reflector structure at which 100% dyke injection (oceanic crust in the strict sense) occurs may well vary regionally: consequently model 5 cannot be ruled out entirely, as it can be regarded as a special case of model 3 in which extension has occurred essentially instantaneously. Results of detailed isotope work on the Voring Plateau sequence of Site 642 are therefore awaited with keen interest.

The authors thank Lindsay Parson, David Roberts, David Smythe and Olav Eldholm for their valuable discussion and comments. This contribution is published with the approval of the Director of the British Geological Survey (NERC).

Received 16 June 1986; accepted 5 January 1987.

- Roberts, D. G., Montadert, L. & Searle, R. C. *Init. Rep. DSDP* 48, 1061-1088 (1979).
- Featherstone, P. S., Bott, M. H. P. & Peacock, J. H. *Geophys. J. R. astr. Soc.* 48, 15-27 (1977).
- Larsen, H. C. in *Petroleum Geology of the North European Margin* 329-339 (Graham & Trotman, London, 1984).
- Smythe, D. K. *Structure and Development of the Greenland-Scotland Ridge* (eds Bott, M. H. P. *et al.*) 109-119 (Plenum, New York, 1983).
- Hinz, K. *Geol. Jb.* E22, 3-28 (1981).
- Mutter, J. C., Talwani, M. & Stoffa, P. L. *Geology* 10, 353-357 (1982).
- Roberts, D. G., Backman, J., Morton, A. C., Murray, J. W. & Keene, J. B. *Init. Rep. DSDP* 81, 883-911 (1984).
- Talwani, M. *et al. Init. Rep. DSDP* 38, 1-1256 (1976).
- Roberts, D. G. *et al. Init. Rep. DSDP* 81, 1-923 (1984).
- Roberts, D. G. & Ginzburg, A. *Nature* 308, 435-437 (1984).
- Smythe, D. K., Chalmers, J. A., Skuce, A. G., Dobinson, A. & Mould, A. S. *Geophys. J. R. astr. Soc.* 72, 373-398 (1983).
- Talwani, M., Mutter, J. C., Eldholm, O. & Stoffa, P. L. *Geophys. J. R. astr. Soc.* 78, 627-637 (1984).
- Garipey, C., Ludden, J. & Brooks, C. *Earth planet. Sci. Lett.* 63, 257-272 (1983).
- Eldholm, O. *et al. Proc. Ocean Drilling Programme* 104A (in the press).
- Harrison, R. K. & Merriman, R. J. *Init. Rep. DSDP* 81, 743-774 (1984).
- Richardson, C., Oakley, P. J. & Cann, J. R. *Init. Rep. DSDP* 81, 795-806 (1984).
- Joron, J. L., Bougault, H., Maury, R. C., Bohn, M. & Desprairies, A. *Init. Rep. DSDP* 81, 783-794 (1984).
- Macintyre, R. M. & Hamilton, P. J. *Init. Rep. DSDP* 81, 775-781 (1984).
- Dickin, A. P. *J. Petrol.* 22, 155-189 (1981).
- Cohen, R. S., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* 283, 149-153 (1980).
- Dupré, B. & Allègre, C. J. *Nature* 286, 17-22 (1980).
- Taylor, S. R. & McLennan, S. M. in *The Continental Crust: its Composition and Evolution* 1-312 (Blackwell Scientific, 1983).
- Miller, J. A., Matthews, D. H. & Roberts, D. G. *Nature phys. Sci.* 246, 61 (1973).

Helium isotope disequilibrium and geochronology of glassy submarine basalts

David W. Graham, William J. Jenkins
Mark D. Kurz & Rodey Batiza*†

Department of Chemistry, Woods Hole Oceanographic Institution,
Woods Hole, Massachusetts 02543, USA

* Department of Earth and Planetary Sciences,
Washington University, St. Louis, Missouri 63130, USA

The (U+Th)/He dating method is the oldest method of geochronology that involves radioactivity, but historically it has not been particularly successful. Even when problems of poor helium retention and loss due to radiation damage are circumvented by analysing suitable mineral phases¹⁻⁴, problems frequently remain with inherited helium^{5,6}. Here we show that isotope disequilibrium in the $^3\text{He}/^4\text{He}$ ratio between helium trapped in vesicles and that dissolved in the glass phase of some young seamount basalts can be used to determine (U+Th)/He ages for the basalts. We use the $^3\text{He}/^4\text{He}$ in vesicles (extracted by crushing *in vacuo*) to correct the dissolved phase He (by fusion of the remaining powder) for the inherited component, and compute the radiogenic helium concentration. The method is applicable to rocks containing phases with different (U+Th)/He, and the results have implications for dating lavas in the age range of 10^3 to 10^6 years (thus filling a gap between the limits of the ^{14}C and K-Ar methods), and for reconstructing the geochemical history of young volcanic systems.

A suite of basaltic glasses from small seamounts near the East Pacific Rise at $10-14^\circ\text{N}$ has been analysed. This seamount field erupts chemically diverse lavas ranging from tholeiite (with mid-ocean-ridge basalt (MORB) chemistry) to alkali basalt (from a source with time-integrated higher Rb/Sr, Nd/Sm and (U+Th)/He⁷⁻⁹). Here we report data for seamount 6, a small volcano (volume = 60 km^3 ; height = $1,300\text{ m}$) located on 3 Myr-old seafloor ($12^\circ44'\text{N}$, $102^\circ35'\text{W}$)¹⁰. This seamount consists of three coalesced volcanoes, which have erupted a broad range of lava types. In all but one case, the alkali-rich lavas show He isotope disequilibrium between vesicles and glass (Table 1; Fig. 1). The glass phase shows a clear radiogenic signature, with $^3\text{He}/^4\text{He}$ much less than the atmospheric ratio ($R_a = 1.4 \times 10^{-6}$). Because the alkalic rocks are both enriched in U and Th and have low He concentrations with respect to tholeiites^{8,9}, radiogenic ingrowth of ^4He is clearly detectable on the short timescale involved, that is 10^3 to 10^6 yr. Transitional basalts, presumably derived by mixing of tholeiitic and alkalic magmas^{7,8}, sometimes show He isotope disequilibrium. Only one of 22 tholeiites from this seamount field shows significant disequilibrium.

Following *in vacuo* crushing and analysis¹¹, the $<100\text{ }\mu\text{m}$ fraction is transferred to a clean aluminium foil boat. The powders are then analysed for the remaining helium by melting *in vacuo* in a resistively heated Ta crucible, contained within a secondary vacuum envelope of a water-cooled furnace¹². Furnace hot blanks are run before and after every sample. The average blank attributable to the Al boats is undetectable. The total processing blank for crushing and melting is the same, currently 4.3 ± 0.3 (one standard deviation) $\times 10^{-11}\text{ cm}^3\text{ STP He}$. Results for three samples of different ages have been replicated, and one sample (6B2) was also analysed by melting whole glass shards. Calculated ages for these cases are in agreement within analytical uncertainty (Table 1).

The minimum detectable age is dependent upon the U+Th content and the detection limit for He isotope disequilibrium.

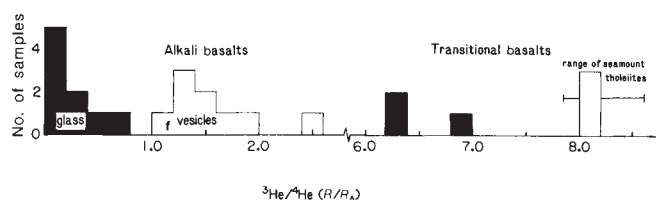


Fig. 1 Histogram of $^3\text{He}/^4\text{He}$ in samples showing isotopic disequilibrium between vesicles and glass at seamount 6. The range of $^3\text{He}/^4\text{He}$ in 22 other tholeiites and transitional basalts from young seamounts near the East Pacific Rise is shown for comparison.

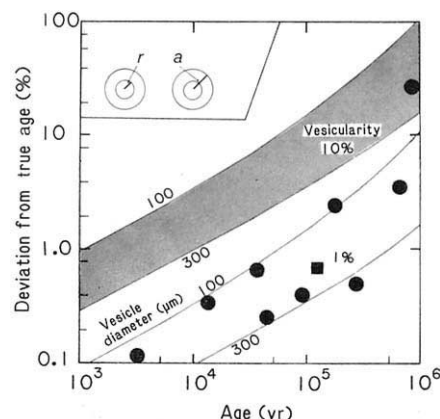


Fig. 2 Model for partial vesicle-glass equilibration by diffusion. The deviation (lowering) of the age T (in %) as a function of the true age is shown for vesicularities of 1% and 10% and a range of vesicle diameters from $100-300\text{ }\mu\text{m}$. The diffusion coefficient, D , of He in basalt glass is $10^{-17}\text{ cm}^2\text{ s}^{-1}$, ref. 11. Vesicle geometry in this model is assumed to be the simplest case of homogeneously distributed spheres of uniform size; vesicle radius is r and diffusive radius $a = r + (DT)^{1/2}$. For low vesicularities ($<10\%$) an individual vesicle domain is defined by a surrounding glass volume, $V = (p)(4\pi/3)r^3/(\text{ves})$, where ves is the volume fraction of vesicles and p is a geometric factor accounting for the vesicle packing fraction. A concentric shell in diffusive communication with a vesicle has a thickness approximated by $(DT)^{1/2}$, and volume $V_D = 4\pi/3\{(r + \sqrt{DT})^3 - r^3\}$. The fractional lowering (f) of the apparent age from the true age will be the fraction of $^4\text{He}^*$ missing from the glass which is lost to vesicles. For glass cubes with uniform U and Th distribution ($p = 8/(4\pi/3)$), $f = V_D/V = (\text{ves})/(\pi/6)\{(r + \sqrt{DT})^3/r^3 - 1\}$. Circles are alkali basalts and the square is a transitional basalt from seamount 6. The diffusive effects are not appreciable in this model except in the case of the oldest sample.

The amount of radiogenic He produced in a time T (for $T < 10^7\text{ yr}$) is

$$^4\text{He}^* = 2.80 \times 10^{-8} \{ [U](4.35 + \text{Th}/U) \} T \text{ (cm}^3\text{ STP g}^{-1}) \quad (1)$$

where T is in Myr, Th/U is the atom ratio, and $[U]$ is the uranium concentration in p.p.m. We use accepted decay constants¹³ and assume secular equilibrium. The limiting factors in detecting ingrown ^4He are the reproducibility of the processing blank for analysis by mass spectrometry, and contamination from inherited He. For most of the present work the latter is the major limitation. The ^3He detection limit ($\sim 1.4 \pm 3 \times 10^{-17}\text{ cm}^3\text{ STP}$) may also become significant in distinguishing small age differences in very young samples. We estimate that the minimum detectable age is $\sim 5,000$ years for the sample types analysed here.

(U+Th)/He ages are calculated using several simplifying assumptions. The glass is assumed to be a closed system since

† Present address: Department of Geological Sciences, Northwestern University, Evanston, Illinois 60201, USA.

Table 1 Chemical analyses of seamount 6 basalts

Sample	Rock type*	Mg no.	(La/Sm) _N	Vesicularity (vol. %)	Wt. (mg)	Crushed				Melted				U (p.p.m.)	Th (p.p.m.)	Age 10 ³ yr
						[He] 10 ⁻⁹ (cm ³ STP g ⁻¹)	³ He/ ⁴ He (R/R _A)	±σ	Wt. (mg)	[He] 10 ⁻⁹ (cm ³ STP g ⁻¹)	³ He/ ⁴ He (R/R _A)	±σ	⁴ He* 10 ⁻⁹ (cm ³ STP g ⁻¹)			
Rise D19-23†	AB	56.1	2.70	6.0	306.1	55.6	1.39	0.02	117.4	220	0.12	0.01	200	1.0	3.6	900±300
Rise D20-2	AB	52.5	1.90	1.0	190.4	9.9	1.39	0.05	93.9	19.9	0.33	0.05	15.2	0.79	2.8	88±10
Ceres D6-3	AB	51.9	2.72	0.5	123.1	3.5	1.86	0.28	96.8	15.0	0.14	0.07	13.9	1.52	5.02	43±3
Ceres D6B2	AB	54.3	2.84	1.5	175.7	10.4	1.58	0.07	70.5	15.7	0.42	0.10	11.6	1.47	5.25	36±4
replicate whole glass shards					369.0	9.4	1.51	0.15	105.0	13.5	0.36	0.10	10.3			32±4
Ceres D6-6‡	TH	65.1	0.28	ND	72.3	154	8.11	0.04	36.2	25.9	6.94	0.20	3.7	0.04	ND	500±500
ALV1388-1724	TR	65	1.13	ND	103.2	60.9	8.13	0.05	44.7	27.9	6.38	0.15	6.0	0.24	0.80	120±25
ALV1389-1613	AB	57.3	2.52	0.2	259.6	2.9	1.48	0.12	92.7	1.2	0.60	0.95	0.7	1.19	3.87	3±5
replicate					403.2	3.3	1.31	0.12	60.6	1.0	3.3	2.0	1			<4
ALV1389-1647	AB	55.0	2.90	ND	239.5	3.2	1.20	0.14	87.6	62.1	0.19	0.02	52.4	1.42	4.79	170±12
ALV1389-1810	AB	61.6	2.68	0.6	241.7	3.7	2.55	0.17	101.4	77.5	0.28	0.02	69.1	1.20	3.84	270±17
replicate					218.1	3.9	3.01	0.4	78.9	73.2	0.33	0.03	65.2			255±16
ALV1389-1854B	AB	54.7	ND	1.0	304.8	6.2	1.67	0.03	112.9	216	0.06	0.01	210	1.47	4.58	680±140
ALV1389-2115B	AB	56.0	2.9	1.3	295.9	6.3	1.32	0.07	118.4	4.5	0.05	0.18	4.3	1.57	5.18	13±3
ALV1391-2033	TR	64.6	1.16	0.4	210.2	57.3	8.17	0.03	89.7	17.7	6.24	0.10	4.2	0.16	0.78	100±33

ND, not determined. Crushed analyses are for 2–4 mm grain size; melted analyses are for the remaining <100 μm powder. U and Th determined by neutron activation analysis (ref. 8); typical uncertainties are 10% and less than 5%, respectively. Uncertainties in the computed ages include those associated with both U, Th and ⁴He* determinations. In most cases, analytical uncertainty for [U] contributes the largest error in the estimated age.

* AB, alkali basalt; TR, transitional basalt; TH, tholeiite.

† Micro-crystalline surface texture. All other samples show vitreous surfaces.

‡ ~10% of the glass in this sample is palagonitized. If the observed disequilibrium is partly due to retention of radiogenic He in palagonite, the computed age may be an upper limit.

§ Calculated from the relative constancy of K/U and Th/U in basalts; K/U = 1.27 × 10⁴ and Th/U = 3.5 (ref. 23); uncertainty in the ratios for these seamount basalts is ~30%.

|| Calculated from measured [U] assuming Th/U = 3.5.

the time of eruption, with no isotope re-equilibration between vesicles and glass. The radiogenic He is

$${}^4\text{He}^* = [\text{He}]_g \left\{ \frac{R_v - R_g}{R_v - R_r} \right\} \quad (2)$$

where g and v refer to glass and vesicle phases and $R = ({}^3\text{He}/{}^4\text{He})$. R_r , the radiogenic production ratio, depends upon sample composition and is between 10⁻⁸ and 10⁻⁷ for basalts¹⁴⁻¹⁶. (⁶Li(n, α)³H → ³He, where the neutron production results from α -bombardment of Mg, Si and O target atoms, is by far the most important reaction producing ³He *in situ* for seafloor basalts). Since $R_r \ll R_v$, approximating equation 2 as

$${}^4\text{He}^* = [\text{He}]_g \{1 - (R_g/R_v)\}$$

does not presently introduce appreciable errors. Combining (1) and (2) gives

$$T = 3.58 \times 10^7 \{ [\text{He}]_g (1 - (R_g/R_v)) / [U] (4.35 + \text{Th}/U) \} \quad (3)$$

Some processes considered in the following discussion which could potentially affect inferred eruption ages include (i) atmospheric contamination; (ii) alteration of U and Th concentrations; (iii) implantation of α -particles from relatively U, Th-rich alteration surfaces and (iv) He loss from the glass.

Vesicularities of the samples range between 0.2 and 6 volume per cent (Table 1), with vesicle diameters between 0.1–0.5 mm. The ratio of ³He trapped in vesicles to that dissolved in the glass is similar to the value determined for MORB at comparable vesicularities, where isotope disequilibrium has not been observed¹¹. The fact that partitioning of inherited He (³He) is similar for these different rock types suggests that no isotope fractionation occurred between gas and melt during eruption. Atmospheric contamination of the dissolved phase He is insignificant, as the ³He/⁴He ratio is much less than 1R_A and close to the theoretically predicted production ratio for basalts. In addition, the dissolved [He] in a basalt at equilibrium with air-saturated sea water is 2 × 10⁻⁹ cm³ STP g⁻¹ (for Henry's law constant $K = 3.7 \times 10^{-4}$ cm³ STP g-atm⁻¹, refs 5, 11, 12), whereas measured concentrations in the alkali basalts range up to 2 ×

10⁻⁷ cm³ STP g⁻¹. At the present time we cannot exclude the possibility that the vesicles of these alkali basalts contain a fraction of an atmospheric (or sea-water) He component, but such contamination requires He migration along cracks because volume diffusion is too slow. We have found no evidence for the presence of significant cracks in either vesicle walls or the surrounding glass using the scanning electron microscope.

U and Th addition to glass could have varying effects on measured ages depending upon the degree of retention of the generated α -particles. Surficial U and Th may implant α -particles, but this process also causes some radiation damage, and most of the added He will likely be lost due to enhanced diffusion. Palagonite and manganese crusts are the most important U and Th-rich alteration phases^{19,20}. We have avoided any significant contribution from alteration phases by careful cleaning and microscopic selection of very fresh samples that show vitreous surfaces on all sides.

Some possible mechanisms of He loss include migration along high diffusivity paths, α -ejection and volume diffusion of He, which might lead to loss from vesicles or glass, or to their partial isotope re-equilibration. Helium loss to high diffusivity paths is related to the population of micro-cracks, which ultimately depends upon the thermo-mechanical response of the glass to cooling and other changes. Significant (>10%) loss requires crack spacing to be ~10 times the diffusive distance for a given sample age, and of appropriate geometry that He loss is effective. For a diffusion coefficient (D_{He}) ≈ 10⁻¹⁷ cm² s⁻¹ at sea-floor temperatures¹¹ and ages of 10⁴ to 10⁵ yr, crack scale lengths of 10 to 100 μm are required. We see no evidence for such cracks with the scanning electron microscope. Ultimately, weathering and devitrification are likely to be more important violations of glass remaining a closed system in (U+Th)/He dating, but careful sample selection avoids these problems.

The average α -stopping distance for a basalt glass is 20 μm. Only 25% of the ⁴He generated within a surface layer of this thickness can be ejected from the sample²¹. Such a loss produces <1% uncertainty in the computed age for a sample from the outer 2 mm of a basalt pillow. Our samples are from pillow rims

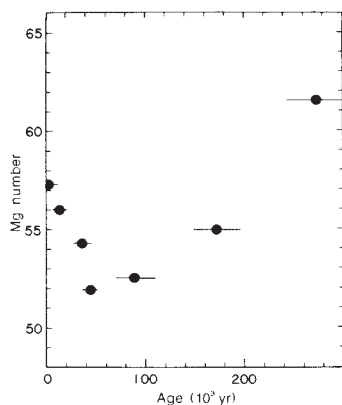


Fig. 3 Mg number ($\equiv 100 \times \text{Mg}/(\text{Mg} + \text{Fe}^{2+})$) against age for the youngest seamount 6 alkali basalts. ± 2 -sigma error estimates in the age are indicated.

typically 0.5–1.0 cm thick. The diffusion coefficient for He in basaltic glass is presently a point of debate^{11,17}, but firm upper limits for the extent of diffusive loss ($D_{\text{He}} \approx 10^{-15} \text{ cm}^2 \text{ s}^{-1}$ at sea-floor temperatures and 0.5 cm pillow rim thickness) are 4% and 11% in 10^4 yr and 10^5 yr, respectively. These losses, although significant, are comparable to our uncertainties.

We can estimate the extent of reequilibration between glass and vesicles using a simple diffusion model that assumes homogeneous distribution of spherical vesicles with uniform radius. Results of this model are shown in Fig. 2 along with the apparent age reduction that would occur for seamount 6 samples. The true age is increasingly underestimated as vesicularity increases and as vesicle size decreases. A typical deviation from the true age for our samples (10^5 yr, 1% vesicularity and 300 μm vesicle radius) is about 1%. Such a calculation is simplistic (for example, it neglects the size distribution of vesicles) but it demonstrates the relative magnitude of the reequilibration effect. Partial reequilibration between vesicles and glass also leads to lower vesicle $^3\text{He}/^4\text{He}$. Although this may be a significant effect ($\geq 10\%$) on our estimate of the inherited $^3\text{He}/^4\text{He}$ for older samples, it has much less effect on the calculated age because the inherited component of the glass phase He is a smaller fraction of the measured [He]. In addition, if more than one vesicle size population is present, the larger vesicles (which are less affected by such reequilibration) will dominate the measured $^3\text{He}/^4\text{He}$ because of larger He contents. We conclude that although there are some inherent complications in the (U+Th)/He technique, the uncertainties introduced are comparable to, or less than, the analytical error for the studied samples.

At seamount 6, a pair of transitional basalts from the eastern satellite cone have similar ages ($\sim 10^5$ yr). Another pair (the two Ceres 6 alkali basalts, dredged from a terrace between this eastern cone and the main vent¹⁰), also have similar ages ($\sim 40,000$ yr). Collectively, the alkali basalts show a smooth trend in Mg number against age, with a minimum between 5×10^4 to 10^5 yr ago (Fig. 3). The approximate agreement between the age of the eastern cone and the age of the Mg number minimum suggests that a change in magmatic 'plumbing' occurred near this time. The event is marked by subsequent supply of less differentiated magma to the central volcanic complex. Although the spatial distribution of ages for the summit lavas (Rise and ALV1389 samples, see Table 1) is complex, they do not appear to violate principles of stratigraphic superposition. Additionally, the uniform ages obtained for the two distinct areas away from the main summit reinforce our contention that the ages are geologically meaningful, representing times of eruption, and thus provide a means to study volcanic evolution. The youth of the exposed lavas at seamount 6 has been confirmed

by Alvin dive observations (for example, the lack of significant sediment cover) and is also consistent with magnetic gradiometer measurements²². All evidence points to volcanism occurring at seamount 6 at a significant distance off-axis of the East Pacific Rise.

We thank Teri Smith for access to unpublished seamount 6 data and for helpful discussions about its geology. We thank Tom Trull, Peter Meyer and Stan Hart for helpful and interesting discussions. We thank Dempsey Lott, whose careful engineering design, construction and modification of the mass spectrometer and sample processing system made the analyses possible. This work was supported by the NSF. D.G. thanks the Education Office of the MIT/WHOI Joint Program in Oceanography for financial support.

Received 21 July; accepted 11 December 1986.

1. Keevil, N. B. *Amer. J. Sci.* **241**, 680–693 (1943).
2. Hurley, P. M. *Geol. Soc. Amer. Bull.* **61**, 1–8 (1950).
3. Fanale, F. P. & Kulp, J. L. *Econ. Geol.* **57**, 735–746 (1962).
4. Damon, P. E. & Green, W. D. in *Proc. on the Symposium on Radioactive Dating* (Int. Atomic Energy Agency, Vienna 1963).
5. Damon, P. E. & Kulp, J. L. *Amer. Mineral.* **43**, 433–459 (1958).
6. Noble, C. S. & Naughton, J. J. *Science* **162**, 265–267 (1968).
7. Zindler, A. *et al. Earth planet. Sci. Lett.* **70**, 175–195 (1984).
8. Batiza, R. & Vanko, D. *J. geophys. Res.* **89**, 11235–11260 (1984).
9. Graham, D. W. *et al. EOS* **65**, 1079 (1984).
10. Batiza, R. & Vanko, D. *Mar. Geol.* **54**, 53–90 (1984).
11. Kurz, M. D. & Jenkins, W. J. *Earth planet. Sci. Lett.* **53**, 41–54 (1982).
12. Kurz, M. D., Jenkins, W. J. & Gurney, J. J. *Earth planet. Sci. Lett.* (in the press).
13. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359–362 (1977).
14. Morrison, P. & Pine, J. *Ann. N.Y. Acad. Sci.* **62**, 69–92 (1955).
15. Mamyrin, B. A. & Tolstikhin, L. N. *Helium Isotopes in Nature* (Elsevier, Amsterdam, 1984).
16. Andrews, J. N. *Chem. Geol.* **49**, 339–351 (1985).
17. Jambon, A., Weber, H. W. & Begemann, F. *Earth planet. Sci. Lett.* **73**, 255–267 (1985).
18. Marty, B. & Ozima, M. *Geochim. cosmochim. Acta* **50**, 1093–1097 (1986).
19. Newman, S. *et al. Earth planet. Sci. Lett.* **65**, 17–33 (1983).
20. MacDougall, J. D. *et al. Earth planet. Sci. Lett.* **42**, 27–34 (1979).
21. Bender, M. L. thesis, Columbia Univ. (1970).
22. McNutt, M. J. *geophys. Res.* **91**, 3686–3700 (1986).
23. Jochum, K. P. *et al. Nature* **306**, 431–436 (1983).

Electroreceptors in the platypus

J. E. Gregory, A. Iggo*, A. K. McIntyre & U. Proske

Department of Physiology, Monash University, Clayton, Victoria 3168, Australia

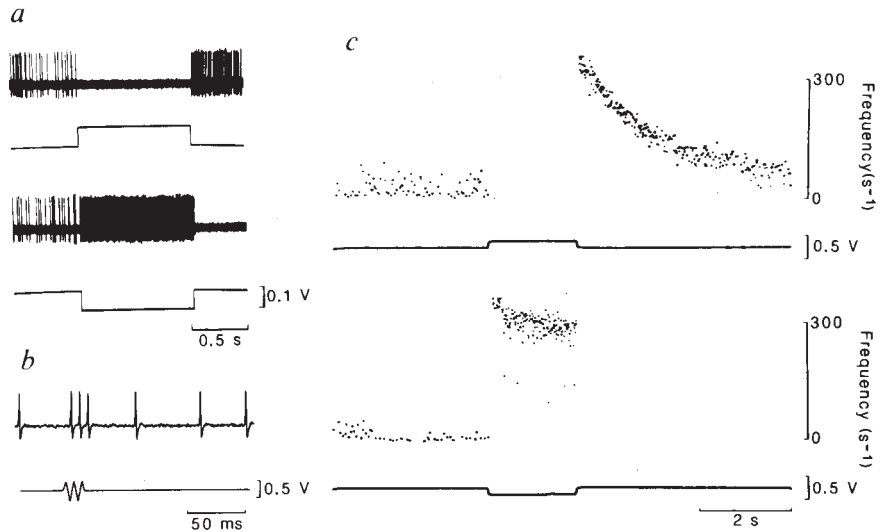
It has been known since the last century that the bill of the platypus contains densely packed arrays of specialized receptor organs and their afferent nerves. Until recently these were thought to be largely mechanoreceptive in function. However Scheich *et al.*¹ provide both behavioural and electrophysiological evidence that there are electroreceptors in the bill of the platypus. These authors were able to record evoked potentials from the somatosensory cortex of the brain in response to weak voltage pulses applied across the bill. Behavioural observations showed that a platypus could detect weak electric dipoles and it was suggested the animal was able to locate moving prey by the electrical activity associated with muscle contractions. From these observations, and in view of the fact that it was known that the bill contained gland receptors² which in several respects resembled the ampullary electroreceptors in freshwater fish, Scheich *et al.* concluded that the receptor array of the platypus bill included electroreceptors. In this report we present direct electrophysiological evidence for the existence of such receptors.

Discharges recorded in the infraorbital nerve which supplies the bill could be identified as coming from either mechanoreceptors or electroreceptors. Electroreceptors could be readily distinguished from mechanoreceptors as the electroreceptors showed

* Permanent address: Department of Preclinical Veterinary Sciences, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, UK.

Fig. 1 Records from electroreceptors in platypus bill, showing responses to electrical stimuli applied between a pair of electrodes in contact with the bill surface. One of the electrodes was placed in the position found to be maximally effective, presumably directly over the receptor. In *a*, an anodal stimulus delivered through this electrode abolished the resting discharge and was followed by rebound excitation (upper pair of traces) whereas cathodal stimulation strongly excited the receptor and was followed by rebound inhibition (lower pair of traces). Scale bars 0.1 V (vertical), 0.5 s (horizontal). *b*, Response of another receptor to a brief train of triangular pulses, shaped to imitate the potentials recorded in the surrounding water during a shrimp's tail flick. Scale bars 0.5 V (vertical), 50 ms (horizontal). In *a* and *b*, in each of the three pairs of traces, action potentials are shown in the upper, and potential difference between the stimulating electrodes in the lower, trace. In *c*, the responses (upper trace in each pair) of a third receptor to anodal, inhibitory (above) and cathodal, excitatory (below) stimuli are shown as instantaneous impulse frequency (vertical axis). Stimulus scale bars: 0.5 V (vertical); 2 s (horizontal).

Methods. Experiments were carried out on one male and one female platypus (*Ornithorhynchus anatinus*) which had been anaesthetized with an intraperitoneal dose of chloralose (60 mg per kg), and were artificially ventilated and maintained at their normal body temperature of 32 °C by means of a heating blanket and rectal probe. The infraorbital branch of the trigeminal nerve was exposed and small strands dissected away from the main trunk for recording of afferent activity on bipolar platinum electrodes. Electrical stimulation of the bill was carried out using a pair of electrodes made of thin silver or platinum wire. The ends of the wires were placed in contact with the bill, which was kept moist by regularly irrigating it with tap water or, on occasions, covering it with a layer of wet cotton wool.



measurable responses to electrical stimuli in the millivolt range and were insensitive to non-noxious probing with a fine glass stylus. Direct electrical stimulation of mechanoreceptors required voltages which were about three orders of magnitude larger. Furthermore the majority of electroreceptors had an irregular resting discharge at 30–50 impulses per second, whereas in the absence of an applied stimulus mechanoreceptors typically were silent. Systematic exploration of a portion of the infraorbital nerve suggested that a majority of the afferents in it supplied electroreceptors.

Placing the stimulating electrodes at different sites over the bill had little effect on electroreceptor threshold. However for each receptor a small spot of maximum sensitivity could be located which, under the dissecting microscope, was seen to coincide with a single pore in the skin. Several of these pores were marked during the experiment with small pins and subsequent histological examination showed them to be the openings of large innervated gland ducts. An example of the response to suprathreshold stimulation at the site of maximum sensitivity is shown in Fig. 1. Stimulating pulses with the cathode directly overlying the sensitive spot and the other electrode elsewhere on the bill produced a large increase in discharge during stimulation, with some rebound inhibition at the end of the stimulus; anodal pulses silenced the discharge and were followed by a large rebound excitation (Fig. 1*a*). When afferent discharge was displayed as instantaneous frequency (Fig. 1*c*), the irregularity in the resting discharge became more apparent, and some adaptation could be detected during the application of the cathodal pulse. Peak rates of discharge during stimulation exceeded 300 impulses per second. Note the very large rebound excitation following anodal stimulation.

The threshold for step-function stimulation of electroreceptors, with the cathode overlying the point of maximum sensitivity, was typically about 20 mV but could become as much as 1 V when the electrode was not directly over the sensitive spot. It is of interest that Scheich *et al.* obtained behavioural evidence for responses to voltage gradients as low as 60 V cm⁻¹. Although our threshold values cannot be directly compared with those of Scheich *et al.*, because of the different methods of stimulation

used, it seems likely that the values obtained for individual receptors represent significantly higher thresholds than those from the behavioural experiments. Temporal and spatial integration of the activity from many receptors may allow the animal to detect much lower voltage gradients.

Electroreceptors responded to rapidly changing voltages. This is of particular interest as Scheich *et al.* had suggested that the electrical activity accompanying a tail flick of the freshwater shrimp could be detected by platypus. We have simulated the waveform recorded by them from a shrimp's tail flick and used it to test the sensitivity of electroreceptors. With sufficiently large pulses a measurable response could always be elicited. For the unit illustrated in Fig. 1, the threshold was 14 mV, and at strengths above the threshold the response was 'locked' to the stimulus (Fig. 1*b*).

The fact that in platypus electroreceptors are supplied by the trigeminal (5th cranial) nerve, not the 8th cranial nerve as in fish, strongly suggests the independent evolution of electroreception in monotremes. The ampullary electroreceptors of sharks and rays respond readily to steady potentials and to low frequency sinusoidal potential fluctuations³. Our data suggest that the platypus, too, is able to detect d.c. fields but its electroreceptors have a much wider dynamic range. The frequency of the voltage fluctuations in Fig. 1*c* was 140 Hz, beyond the detection limits of the elasmobranch receptors³. Platypus are common in muddy streams and at most times have eyes, ears and nose shut during a dive, so they probably rely almost entirely on electro-detection to locate prey, and to receive information about obstacles on the stream bottom, and variation in the depth of the stream.

Received 16 October 1986; accepted 4 February 1987.

1. Scheich, H., Langner, G., Tidemann, C., Coles, R. B. & Guppy, A. *Nature* **319**, 401–402 (1986).
2. Andres, K. H. & Von Düring, M. in *Sensory Receptor Mechanisms* (eds Hamann, W. & Iggo, A.) 81–89 (World Science, Singapore, 1984).
3. Kalmijn, A. J. in *Handbook of Sensory Physiology* Vol. III/3 (ed. Fessard, A.) 147–200 (Springer, Berlin, 1974).

Habitat fragmentation and the stability of predator-prey interactions

P. Kareiva

Department of Zoology NJ-15, University of Washington, Seattle, Washington 98195, USA

Mathematical models¹⁻³, field observations^{4,5}, and laboratory studies⁶ all suggest that habitat patchiness (or 'fragmentation') profoundly affects species interactions. One especially widely cited idea is that patchiness stabilizes predator-prey dynamics^{7,8}. I performed the first test of this idea in a natural community by experimentally manipulating the degree of patchiness in goldenrod fields that were the setting for a predator-prey interaction between ladybird beetles and aphids. Contrary to conventional wisdom, I found that increasing patchiness led to more frequent local explosions of aphid populations and thus less stable dynamics. These results can be understood by examining the effects of patchiness on the searching and aggregation behaviour of ladybird predators. It appears that the effects of habitat fragmentation depend on the specific behaviour of the organisms using the habitats. Thus, instead of making robust generalizations about habitat fragmentation (such as "patchiness is stabilizing") we should seek predictions that are based on the details of an organism's dispersal behaviour and demography⁹.

My experiments were performed on an abandoned cattle ranch overlooking Newport Bay in Bristol County, Rhode Island. To minimize extraneous sources of variation, I selected fields that were effectively monocultures of the goldenrod, *Solidago canadensis*. The prey species, *Uroleucon nigrotuberculatum*, is an aphid that feeds only on *Solidago* and overwinters at the base of its host plants. Although the predator, *Coccinella septempunctata*, is known to consume a wide variety of aphid species voraciously, at my study site it ate almost exclusively *Uroleucon*¹⁰. The key to my experiment was the manipulation of habitat patchiness by mowing fields into variously shaped pure stands of goldenrod.

In particular, I established paired fragmentation treatments, of 'patchy' or 'continuous' goldenrod surrounded by grass lawn, which represented hostile territory as far as *Coccinella* and *Uroleucon* are concerned (Fig. 1). This experiment was replicated in three different fields, all of which were sampled biweekly during the summers of 1982-1985 (details of sampling methods are given in Fig. 1, legend). Additional replicates were used for short-term demographic and behavioural studies.

Censuses of my experimental plots revealed that both *Uroleucon* and *Coccinella* were consistently more abundant in the patchy than in the continuous goldenrod habitat (Fig. 2 and Table 1). The increase in aphid density in patchy habitats was often enormous. For example, during 1982 and 1984 aphids were ten times more abundant in patchy than in continuous habitats in all three replicates (see Fig. 2). Closer inspection of the census data suggested that the higher aphid densities in the patchy treatment were due to more frequent localized outbreaks of aphids rather than an evenly spread increase in aphid abundance. To support this impression quantitatively, I defined an 'outbreak' as any situation in which aphids attained densities >100 individuals per stem within a metre-square quadrat or patch. During the course of my four-year study, I observed 26 such outbreaks in the patchy goldenrod but only 9 in the continuous goldenrod (in both cases this is out of a possible 120 outbreaks, representing ten metre-squares sampled in each of three replicates over four years).

There was no evidence to suggest that aphid outbreaks were a direct result of habitat patchiness (or any sort of relationship between *Uroleucon* and *Solidago* dispersion). Not only were aphid colonization rates the same in both the patchy and con-

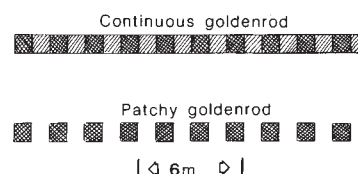


Fig. 1 One replicate of the goldenrod fragmentation experiment. Crosshatched and double crosshatched areas represent goldenrod monocultures surrounded by mown grass. Double-crosshatched quadrats were repeatedly censused at biweekly intervals during the summers of 1982-1985. Single crosshatched areas represent goldenrod that is part of the experimental manipulation (that is, it is part of the continuous treatment), but is not censused. Censuses were performed by counting all aphids and ladybirds on 10-30 stems in each double-crosshatched square-metre. In any given census, the number of stems surveyed was identical in all replicates, and in both treatments; however this number varied between censuses to accommodate seasonally changing frequencies of aphids. Although *Coccinella* could move between the continuous and patchy arrays, most within-field (or within-replicate) ladybird movement was directed along the linear axis of the goldenrod strips or rows of patches. Aphids moved little in any direction.

tinuous treatments (Table 2, methods given in the legend), but, in the absence of predators, net population growth was also the same in both treatments (Fig. 3, methods given in legend). As patchiness did not appear to have any direct effect on aphid abundances, I hypothesized that the differences evident in Fig. 2 were due to the influence of patchiness on the predator-prey interaction between *Uroleucon* and *Coccinella*. I based this hypothesis on published information concerning *Coccinella* and *Uroleucon* as well as on data reported in this paper. For instance,

Table 1 Effect of habitat fragmentation on aphid prey (*U. nigrotuberculatum*) and their ladybird predators (*C. septempunctata*)

Fragmentation significantly increases aphid density (Fig. 2)

Fragmentation significantly increases ladybird beetle density

maximum density in fragmented goldenrod:
3.5 individuals per 10 plants*

maximum density in continuous goldenrod:
2.2 individuals per 10 plants

Fragmentation has no effect on spring colonization rates for aphids

The number of goldenrod stems colonized by at least one aphid by mid-June is:

98 of 1,800 in continuous goldenrod†
84 of 1,800 in fragmented goldenrod

Fragmentation significantly increases the fraction of experimentally created incipient aphid outbreaks

	Fragmented	Continuous‡
Consumed	30%	67%
Persistent	30%	20%
Rapid growth	40%	13%

* Averages from three replicates censused in three different years (1981, 1982, 1984 and 1985 as in Fig. 1); thus $n=9$ for each average. A significant effect was detected by a repeated measures analysis of variance (ANOVA). This test yielded an F -statistic of 8.15, degrees of freedom = 1, 17 and $P < 0.05$.

† No significant difference in these fractions according to a G -test with 1 degree of freedom, $P > 0.25$.

‡ Colonies of 20 aphids were placed on washed golden stems and examined 3 days later. The fate of each experimental colony was assigned to one of three categories: Consumed (<5 aphids), Persistence (6-30 aphids) or Rapid growth (>30 aphids). The percentage of experimental colonies in each of these categories is given for fragmented and continuous goldenrod habitat ($n=30$ for each percentage).

A G -test reveals a significant association between habitat type (fragmented versus continuous) and the fate of the experimental aphid colonies.

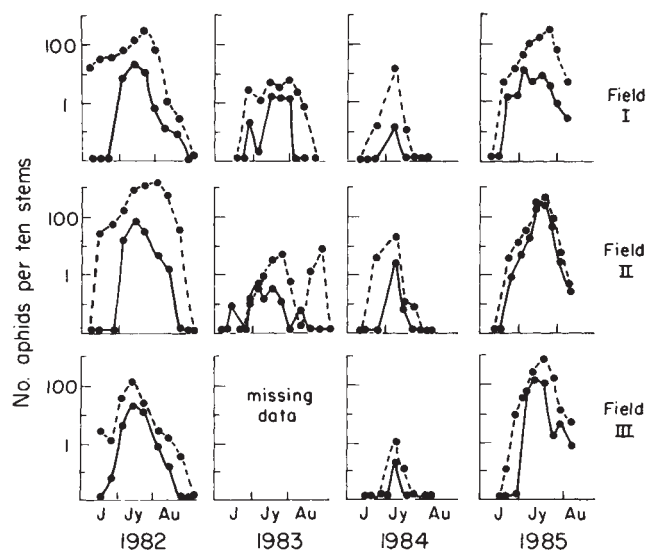


Fig. 2 The influence of goldenrod fragmentation on mean *Uroleucon* density. Aphid densities are plotted on a logarithmic scale, with the dashed lines corresponding to densities in patchy goldenrod and the solid lines corresponding to densities in continuous goldenrod. Each point on these graphs represents an average obtained from census of ten square-metre quadrats, with 10–30 stems inspected per quadrat (thus, between 100–300 goldenrod stems were censused for each data point). To evaluate the statistical significance of the effect of patchiness, I used a repeated-measures ANOVA. I simplified matters by examining only the maximum density the aphids attained in each treatment (the peak in each year's population trajectory). To create a balanced design, the 1983 data were dropped from the analysis, leading to an analysis involving three replicates of two treatments, each with three repeated measures. The results of the ANOVA indicate a highly significant effect of habitat patchiness on maximum aphid density: $F_{1,17} = 20.1$; $P < 0.001$.

when *Coccinella* has been removed from goldenrod plots of any geometry, *Uroleucon* has been consistently shown to increase in density relative to control plots; this indicates that *Coccinella* is at least capable of controlling *Uroleucon* population growth¹¹. A detailed mathematical model based on behavioural observations of *Coccinella* suggests that the ability of these ladybirds to control *Uroleucon* depends on the rapid aggregation of *Coccinella* to clusters of aphids¹². Furthermore, several less specific mathematical models predict that predator aggregation may in general be crucial to the stable control of prey populations^{8,9,13}. Finally, releases of marked *Coccinella* in goldenrod plots similar to those shown in Fig. 1 have revealed that habitat patchiness slows the rate at which ladybirds aggregate at patches of aphids¹⁰.

Taken together, these observations and models lead to the hypothesis that habitat patchiness indirectly promotes aphid outbreaks through its influence on *Coccinella* foraging behaviour. In particular, it seems that *Uroleucon* is usually held at low densities by *Coccinella* because these ladybirds rapidly aggregate at, and exterminate, aphid clusters which threaten to initiate a population outbreak. Habitat patchiness promotes aphid outbreaks because it interferes with the aggregation behaviour of *Coccinella*. Of course, even in patchy habitats ladybirds do eventually aggregate at aphid outbreaks, but once aphids have achieved a threshold density, their reproduction can swamp ladybird predation and aphid population growth becomes limited by different factors, such as the quality of resources provided by the host plant.

I directly tested the above hypothesis by experimentally establishing numerous incipient aphid outbreaks. I placed clusters of 20 aphids on individual goldenrod stems in both patchy and continuous habitats. Three days later I checked these clusters (which had been marked with small coloured tags) to determine

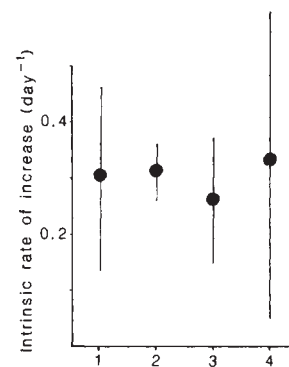


Fig. 3 The influence of goldenrod fragmentation on aphid population growth in the absence of ladybirds. Goldenrod was mowed into patchy and continuous strips as shown in Fig. 1. After colonies had naturally appeared in these strips I transferred them to new stems in one of the following four ways: (1) aphids were transferred from a stem in a continuous goldenrod strip to a different stem in the same square metre of goldenrod vegetation (continuous to continuous); (2) aphids were transferred from a stem in a continuous goldenrod strip to a patch of goldenrod in the fragmented goldenrod habitat (continuous to patchy); (3) aphids from a stem in the fragmented habitat were transferred to a different stem in the same patch of goldenrod (patchy to patchy); (4) aphids were transferred from a stem in the fragmented habitat to a stem in the continuous habitat (patchy to continuous). After the aphid colonies had been transferred, they were checked twice daily and any predators that were found were removed. I recorded the initial number of aphids and the number after 11 days for each experimental colony. To determine the effect of habitat fragmentation on population growth I calculated each colony's intrinsic rate of increase using the formula: $r = \ln(N_{11}/N_0)/11$, where N_{11} was the colony size after 11 days and N_0 was the initial colony size. This model assumes exponential growth, which is appropriate for the densities of aphids used in this experiment. If a logistic model of growth is used the same conclusion is obtained: habitat fragmentation does not influence the rate of population growth when predators are removed. There was no difference in the initial colony sizes between the different treatments.

whether the aphid colonies had persisted or had been eliminated by ladybirds. Because ladybirds leave behind the legs of an aphid after eating it, a ladybird attack provides unmistakable physical evidence (if the carnage is inspected before wind and rain take their toll). I first washed all non-experimental aphids off the goldenrod plants in my study plots to ensure that background aphid densities were equivalent among the two habitat types. I also made certain that ladybird densities were equivalent among the two habitat types. I found that experimental aphid colonies were eliminated by ladybirds twice as frequently in the continuous as in the patchy habitat (Table 1).

In summary, habitat fragmentation appears to promote aphid outbreaks because it interferes with the nonrandom searching behaviour of ladybird predators. Although this result contradicts much prior speculation concerning patchiness and the stability of predator-prey interactions, it confirms ideas about the importance of aggregation behaviour in predator-prey dynamics^{8,9,13,14}. More generally, it is apparent that the influence of habitat fragmentation on any species will be determined largely by that organism's behavioural responses to fragmentation. It is doubtful whether generalizations can be made about the effects of habitat patchiness without detailed behavioural studies. We need such behavioural studies because mankind is exaggerating habitat fragmentation throughout the world and the consequences of this enormous habitat modification are unknown^{15,16}.

I thank the National Science Foundation (NSF DEB 8500343) for financial support and H. Surowiec and R. T. Paine for helpful discussions and editorial comments.

Received 31 December 1986; accepted 22 January 1987.

- Hastings, A. *Theor. Populat. Biology* **12**, 37-48 (1978).
- Levin, S. A. *Rev. ecol. Syst.* **7**, 287-310 (1976).
- Maynard Smith, J. in *Models in Ecology*, 69-83 (Cambridge University Press, 1974).
- Elton, C. J. *Ecol.* **37**, 1-23 (1949).
- Hanski, I. & Ranta, E. *J. Anim. Ecol.* **52**, 263-279 (1983).
- Huffaker, C. B. *Hilgardia* **27**, 343-383 (1958).
- Begon, M., Harper, J. & Townsend, C. in *Ecology*, 337-339 (Sinauer, Massachusetts, 1986).
- May, R. J. *J. Anim. Ecol.* **47**, 833-843 (1978).
- Hassell, M. & Ma, R. in *Behavioural Ecology* (eds Sibly, R. & Smith, R.) 137-154 (Blackwell, Oxford, 1985).
- Kareiva, P. *Lect. Notes Biomath.* **54**, 368-389 (1984).
- Kareiva, P. in *Community Ecology* (eds Diamond, J. & Case, T.) 196-206 (Harper & Row, New York, 1985).
- Kareiva, P. & Odell, G. *Am. Nat.* (in the press).
- Hassell, M. & Ma, R. *J. Anim. Ecol.* **43**, 567-594 (1974).
- Heads, P. & Lawton, J. *Oikos* **40**, 267-276 (1981).
- Wilcox, D., McClellan, C. & Dobson, A. in *Conservation Biology* (ed. Soule, M.) 237-256 (Sinauer, Massachusetts, 1986).
- Pickett, S. & Thompson, J. *Biol. Cons.* **13**, 27-37 (1978).

Changes in abundance or structure of the *per* gene product can alter periodicity of the *Drosophila* clock

Mary Kilar Baylies, Thaddeus A. Bargiello,
F. Rob Jackson* & Michael W. Young

The Rockefeller University, 1230 York Avenue, New York,
New York 10021, USA

* Worcester Foundation for Experimental Biology, Shrewsbury,
Massachusetts 0154, USA

The *period* (*per*) locus, which controls biological rhythms in *Drosophila*, was originally defined by three chemically induced mutations¹. Flies carrying the *per^o* mutation were arrhythmic, whereas *per^s* and *per^l* mutants had circadian behavioural rhythms with 19-hour and 29-hour periodicities, respectively¹. Wild-type flies have 24-hour rhythms. Here we compare the *per* locus DNA sequences of the three mutants with the parental wild-type. The *per^s* and *per^l* mutations lead to amino-acid substitutions, whereas *per^o* introduces an early translation stop (amber). The results indicate that the protein product of *per* controls biological rhythms. We also report that the abundance of this protein may set the pace of the *Drosophila* clock. Although circadian rhythms are restored when arrhythmic (*per^o*) *Drosophila* are transformed with *per* locus DNA, flies receiving identical transforming DNA segments can produce rhythms with periods that differ by more than 12 hours. Transcription studies reveal a tenfold variation in the level of *per* RNA among transformed lines. Levels of *per* RNA are inversely correlated with period length, so that flies with lowest levels of the *per* product have slow-running biological clocks. On the basis of the combined studies we suggest that *per^l* and *per^s* mutants produce hypoactive and hyperactive *per* proteins, respectively.

A number of rhythmic strains were developed from arrhythmic (*per^o*)¹ *Drosophila* by P element-mediated transformation with *per⁺* DNA. All lines received the same *Hind*III fragment of ~7 kilobases (kb) previously shown to contain DNA coding for the 4.5-kb *per* transcript²⁻⁶. The structure of the transforming DNA, designated pCP1, the strategy for selecting transformants, and the DNA sequence of the 7-kb *Hind*III fragment containing the *per* transcription unit have been presented elsewhere^{5,6,23}. All lines were maintained as heterozygous stocks with respect to the transforming DNA as described⁶.

From a preliminary behavioural analysis of 30 transformed lines, it was evident that flies from the different lines could have rhythms with different periodicities. The period lengths ranged from about 25 h to 40 h. For comparison, wild-type, Canton-S flies (from which pCP1 DNA is derived), have rhythms with a 24-h period.

Seven lines with varied periodicities were chosen for a more complete behavioural and molecular analysis. Hybridization of

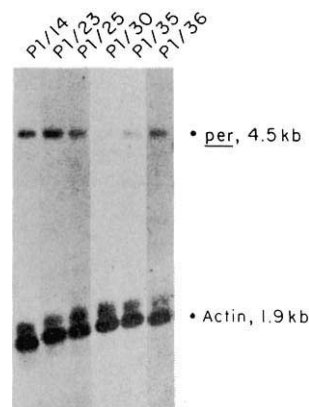
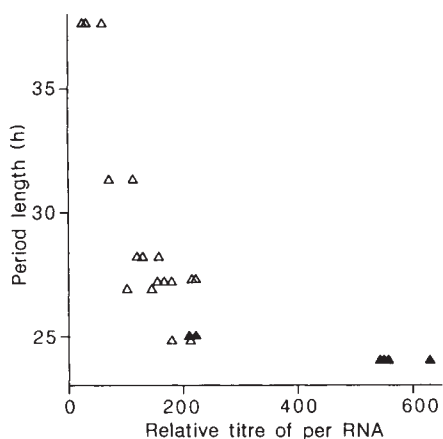


Fig. 1 Abundance of 4.5-kb *per* RNA in different lines of transformed flies. Poly(A)⁺ RNA was isolated from *Df(1)TEM202/Df(1)64j4* adult females heterozygous for transforming DNA of the indicated lines as described⁶. RNA (5 µg per lane) was electrophoresed through 1% agarose/2.2 M formaldehyde gels and transferred to Biotodyne A membrane (Pall). A portion of the *per* transcription unit (*Hind*III-*Eco*RI fragment from coordinates 0.0 to 1.5, ref. 5), was cloned in the Riboprobe vector pSP65 (Promega Biotec). Single-stranded, ³²P-labelled RNA probes were transcribed from the cloned DNA and hybridized to the membrane-bound *Drosophila* RNA at 60 °C in 50% formamide, 50 mM NaPO₄ (pH 7), 0.8 M NaCl, 1 mM EDTA, 0.1% SDS, 1× Denhardt's solution, 250 µg ml⁻¹ denatured salmon sperm DNA, 500 µg ml⁻¹ yeast RNA and 10 µg ml⁻¹ poly(A). The membrane was washed at 65 °C in 5 mM NaCl, 2 mM NaPO₄, pH 7, 1 mM EDTA, 0.1% SDS. A ³²P-labelled actin DNA probe (pDmA2) was also hybridized², as a control for the amount of *Drosophila* RNA loaded on the membrane¹⁷.

cloned *per* locus DNA to polytene chromosomes *in situ* showed that each of the lines carried a single copy of the transforming DNA (see also Table 1). Table 1 shows estimates of period length from analysis of at least 15 flies from each of the seven lines. All period length estimates were from locomotor activity records of adult flies, as described⁴. Again, wide variations in period lengths were observed between lines, but as can be seen from standard error values in Table 1, little variation in periodicity was seen within a line. An exception was strain P1/30 which had a mean period of 37 h. Many of the flies in this strain were arrhythmic (in analysing 15 rhythmic flies, 14 arrhythmic flies were encountered, Table 1), and some rhythmic individuals gave long-period rhythms approaching 45 h. This behaviour is similar to that of very long period, regulatory mutants of *per^s*^{5,7}. The remaining lines showed high penetrance of the rhythmic phenotypes (70-95%, data not shown).

A 4.5-kb poly(A)⁺ RNA is transcribed from the *per* locus in wild-type flies, so comparable transcripts were sought in RNA preparations from transformed lines. Analysis of *per* RNA required a derivative strain from each of the transformed lines. Earlier work has shown that *Df(1)TEM202/Df(1)64j4* females are arrhythmic and homozygously deficient for a 10-kb, X-chromosomal interval that includes the *per* locus²⁻⁴. Thus, strains were crossed to produce *Df(1)TEM202/Df(1)64j4* females carrying autosome-linked transforming DNA. These can produce a *per* transcript only if the transforming DNA is expressed. Transcript analysis of several of the derivative lines is shown in Fig. 1. For this and other experiments (Fig. 2, legend), relative levels of *per* locus transcription were assessed by comparing the abundance of the 4.5-kb *per* RNA to levels of endogenous actin mRNA. All transformed lines produced the 4.5-kb *per* transcript, but the level of accumulated transcript varied among the lines (Fig. 1).

In Fig. 2, levels of 4.5-kb RNA produced by the seven transformed lines are plotted as a function of the period lengths of the lines. An inverse correlation is seen for period length and *per* RNA titre (see also Fig. 2, legend). Also shown in Fig. 2 are



levels of *per* RNA produced by homozygous wild-type *D. melanogaster* females, and females containing only a single copy of *per*⁺ (*per*⁺/*per*^{def}; see Fig. 2, legend). Earlier genetic studies indicated that the latter flies produce circadian rhythms with periodicities slightly longer than wild-type^{8,9}. Figure 2 shows that such flies produce a level of 4.5-kb *per* RNA similar to that of our shortest-period transformed lines. Both kinds of strains have rhythms with about a 25-h periodicity (Table 1 and Fig. 2).

The periodicities of the rhythms of the transformed lines overlap with those of flies carrying wild-type X chromosomes, so it is likely that changes in abundance of the *per* transcript are solely responsible for the period length differences observed in all of the *Drosophila* lines studied. It is not likely that the level of 4.5-kb RNA is itself modulated as a result of changes in period length. If this were the case, flies carrying ethylmethane sulphonate (EMS)-induced *per* locus mutations that shorten or lengthen period should produce higher or lower levels of 4.5-kb RNA. In fact, 4.5-kb transcripts are produced at wild-type levels in these mutants (refs 2, 4 and T.A.B., unpublished results). Presumably variations in *per* RNA titre among transformed lines reflect chromosomal position effects on *per* transcription. Com-

parable position effects were reported in earlier transformation studies of unrelated *Drosophila* genes^{10,11}. The simplest conclusion that can be drawn from these studies is that period length of *Drosophila* circadian rhythms is set by the level of the *per* locus product.

The period of wild-type *Drosophila* rhythms shows long-term stability under constant environmental conditions, adjusts in a limited way to exotic cycles of light and dark, and is relatively insensitive to changes in ambient temperature^{12,13}. Figure 2 shows that it may be possible to view this stability at the molecular level. Throughout the range of periodicities studied, large proportional changes in RNA titre are required for small proportional changes in period length. For example, a threefold drop in RNA titre increases period length from 24 to 27 h, and another threefold drop will lengthen period to about 31 h (Fig. 2). Thus, both changes in RNA level only produce about a 15% change in periodicity.

Period length changes are also associated with chemically induced mutations of the *per* locus. The mutations *per*^s, *per*^o and *per*ⁱ arose, in a single study, following EMS treatment of a *D. melanogaster*, Canton-S stock maintained at the California

Fig. 3 Above, locations of *per* locus point mutations. The structure of the *Eco*RI–*Hind*III fragment sequenced in each mutant is shown. This DNA contains all protein-coding regions of *per* (ref. 5); regions of RNA homology detected by combined heteroduplex⁵ and cDNA^{5,23} studies are indicated by filled boxes. Restriction fragments for sequencing were cloned in M13mp18 or mp19, and ordered sets of overlapping deletions were generated by the method of Dale *et al.*¹⁸. DNA was propagated and sequenced as described^{19,20}. Nucleotide positions for predicted translation start and stop codons in the wild-type gene are given, and locations of single base-pair changes found in *per*ⁱ, *per*^o and *per*^s are shown. Nucleotide numbers are the same as those used to present the Canton-S *per* locus sequence⁵. The open box marks the position of alternating Gly–Thr and Ser–Gly codons at *per*^{5,21,22}. R, *Eco*RI; X, *Xba*I; S, *Sal*I; B, *Bam*HI; H, *Hind*III. Below, *per* DNA and protein sequences in the mutants. Regions of the Canton-S wild-type DNA and protein altered in *per*ⁱ (upper), *per*^o (centre) and *per*^s (lower) are presented. For each mutant, the predicted change in the protein sequence is indicated above, and the change in the DNA sequence below, the corresponding wild-type sequence. Nucleotide numbers are from ref. 5.

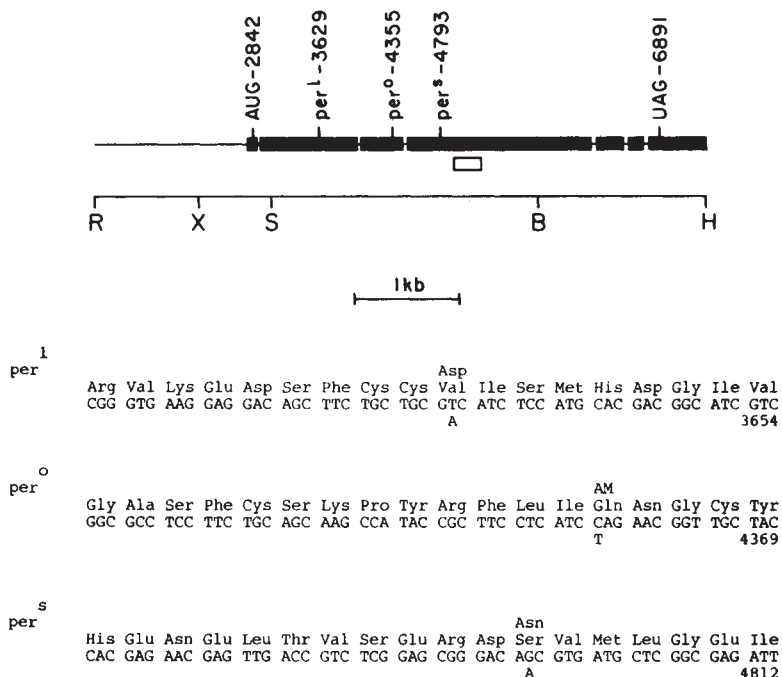


Table 1 Periodicities of transformed *D. melanogaster* lines

Line	Period (h)	No. flies tested	Location of transforming DNA
P1/14	26.9 ± 1.1	20	47E-F
P1/23	27.3 ± 2.6	20	34E-F
P1/25	28.2 ± 1.8	20	97E-F
P1/29	24.8 ± 1.7	21	ND
P1/30	37.6 ± 11.1	15	54A-B
P1/35	31.3 ± 2.0	18	87A-B
P1/36	27.2 ± 1.2	17	93A

See text for methods. Location of transforming DNA refers to polytene chromosome bands: ND, not determined.

Institute of Technology¹. The DNA sequence of a wild-type *per* locus from the parental Canton-S stock was reported earlier⁵. Because all three mutants were derived from the same stock, it is possible to compare DNA sequences among the mutants, as well as with the original wild-type chromosome, to locate changes in the DNA responsible for the altered phenotypes.

Although transformation experiments have shown that *per* corresponds to a ~7-kb transcription unit^{6,14,15}, two studies have also shown that DNA containing a subsegment of the gene (6-kb *EcoRI*-*HindIII* fragment, see Fig. 3 map) is equally capable of restoring rhythms with periodicities in the wild-type range^{14,15}. Sequence analysis of *per* showed that all protein-coding regions of the gene are contained in the latter DNA (refs 5 and 23, and Fig. 3). No change in the abundance or processing of the 4.5-kb *per* transcript was detected in *per*^o, *per*^l or *per*^s mutants (ref. 2; T.A.B. unpublished). Therefore we searched for changes in the sequence of the biologically active *EcoRI*-*HindIII* DNA fragment in the mutants.

A complete DNA sequence was obtained for the interval in all three mutants, yet each gene contained only a single nucleotide substitution compared to the wild-type Canton S sequence. Each nucleotide change was found in only one of the mutants, as would be predicted by their common origin from the parental strain. Thus, each change in the DNA sequence arose in conjunction with the appearance of a mutant phenotype. As shown in the map of Fig. 3 (top), the three mutations affected the structures of different RNA coding regions.

Figure 3 (bottom) shows that in *per*^l and *per*^s DNA, nucleotide substitutions led to changes in the protein sequence, but not in its size. In *per*^l, a T-to-A transversion caused a valine residue to be replaced by aspartic acid. For *per*^s, serine was replaced by asparagine in a G-to-A transition. In *per*^o, a C-to-T transition formed an amber codon (Fig. 3). Thus, *per*^o mutants are expected to produce a protein of ~400 amino acids rather than the wild-type protein of ~1,200 amino acids^{5,23}.

A deletion of the *per* locus produces an arrhythmic phenotype^{2,7,16}, so it is not surprising to find that arrhythmic flies carrying the point mutation *per*^o make a truncated *per* protein. Because *per*^o flies continue to produce wild-type levels of the 4.5-kb RNA^{2,4}, *per* must control biological rhythms through the action of the encoded protein. It is less clear how large changes in period length are effected by *per*^s and *per*^l. From the transformation experiments described in Figs 1 and 2 it can be concluded that abundance of the *per* locus transcript is inversely correlated with the period length of circadian rhythms. If these differences in RNA levels reflect comparable differences in *per* protein titres, then over-production of wild-type protein should result in a phenotype similar to that observed for *per*^s flies, and under-production in a *per*^l-like phenotype. But *per*^s and *per*^l both produce wild-type levels of *per* mRNA (ref. 2 and T.A.B., unpublished), suggesting that in *per*^s and *per*^l, amino-acid substitutions respectively increase or decrease the stability of the *per* protein; or that *per*^s protein has greater

than wild-type activity, whereas *per*^l protein functions subnormally.

We thank Laural Newby-Kew for technical assistance, and Laurel Eckhardt, Peter Model and Norton D. Zinder for comments on the manuscript. This work was supported by grants from the National Institutes of Health and the Andre and Bella Meyer Foundation to M.W.Y.

Note added in proof: two of the point mutations mapped in this study, *per*^o and *per*^s, are also mapped with identical results by Yu *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Received 24 September 1986; accepted 12 January 1987.

1. Konopka, R. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2112-2116 (1971).
2. Bargiello, T. A. & Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2142-2146 (1984).
3. Reddy, P. *et al.* *Cell* **38**, 701-710 (1984).
4. Young, M. W., Jackson, F. R., Shin, H. S. & Bargiello, T. A. *Cold Spring Harb. Symp. quant. Biol.* **50**, 865-875 (1985).
5. Jackson, F. R., Bargiello, T. A., Yun, S.-H. & Young, M. W. *Nature* **320**, 185-188 (1986).
6. Bargiello, T. A., Jackson, F. R. & Young, M. W. *Nature* **312**, 752-754 (1984).
7. Smith, R. F. & Konopka, R. J. *Molec. gen. Genet.* **183**, 243-251 (1981).
8. Smith, R. F. & Konopka, R. J. *Molec. gen. Genet.* **185**, 30-36 (1982).
9. Cote, G. & Brody, S. *J. theor. Biol.* **121**, 487-503 (1986).
10. Spradling, A. C. & Rubin, G. M. *Cell* **34**, 47-57 (1983).
11. Goldberg, D. A., Posakony, J. W. & Maniatis, T. *Cell* **34**, 59-73 (1983).
12. Pittendrigh, C. S. *Proc. natn. Acad. Sci. U.S.A.* **40**, 1018-1029 (1954).
13. Pittendrigh, C. S. in *Handbook of Behavioral Neurobiology* Vol. 4 (ed. Aschoff, J.) 95-124 (Plenum, New York, 1981).
14. Zehring, W. A. *et al.* *Cell* **39**, 369-376 (1984).
15. Hamblen, M. *et al.* *J. Neurogenet.* **3**, 249-291 (1986).
16. Young, M. W. & Judd, B. H. *Genetics* **88**, 723-742 (1978).
17. Fyrberg, E. A., Mahaffey, J. W., Bond, B. J. & Davidson, N. *Cell* **33**, 115-123 (1983).
18. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31-40 (1985).
19. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
20. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
21. Shin, H. S., Bargiello, T. A., Clark, B. T., Jackson, F. R. & Young, M. W. *Nature* **317**, 445-448 (1985).
22. Reddy, P., Jacquier, A. C., Abovich, N., Peterson, G. & Rosbash, M. *Cell* **46**, 53-61 (1986).
23. Citri, Y. *et al.* *Nature* **326**, 42-47 (1987).

DNA marker analysis detects multiple maternity and paternity in single broods of the lesser snow goose

Thomas W. Quinn, James S. Quinn, Fred Cooke & Bradley N. White*

Department of Biology, Queen's University, Kingston, Ontario, Canada K7L 3N6

The determination of the degree of genetic relatedness among individuals is important for many population and behavioural studies¹. Conventional genetic markers such as allozymes are usually inadequate for such determinations, especially in avian species^{2,3}. Restriction fragment length polymorphisms (RFLPs) which are inherited as co-dominant mendelian alleles⁴, have recently been applied to paternity and maternity analyses in humans to address both medical diagnostic and legal⁵ issues. Field observations have led to the suggestion that lesser snow geese (*Anser caerulescens caerulescens*) engage in intraspecific brood parasitism (IBP)⁶ and extra-pair fertilization (EPF)⁷. Direct evidence for these suggestions has been difficult to obtain because of the lack of easily detectable genetic markers⁸. We have recently prepared a recombinant genomic library from the lesser snow goose and isolated seventeen probes that all identify RFLPs⁹. Here we use some of these polymorphic loci to examine the relatedness among members of four families. This study provides evidence for mixed maternity and paternity in broods of this socially monogamous species and demonstrates the utility of DNA markers for genetic analyses of species which traditionally exhibit few conventional genetic markers.

* To whom correspondence should be addressed.

Table 1 Genotypes of nest attendants and goslings of four lesser snow goose families

	Family 1						Family 2					
	♂P1	♀P1	G1	G2	G3	G4	♂P2	♀P2	G5*	G6	G7	G8
DNA marker												
<i>DQSG1-AI</i>	2,2	2,2	2,2	2,2	2,2	2,3†	1,1	1,3	1,3	1,1	1,3	1,3
<i>DQSG1-AII</i>	2,2	2,2	2,2	2,2	2,2	2,2	2,2	2,2	2,2	2,2	2,2	2,2
<i>DQSG1-CI</i>	2,3	2,2	2,2	2,2	2,2	1,1	2,3	1,2	2,2	2,2	1,3	1,2
<i>DQSG5-BI</i>	1,2	1,1	1,2	1,1	1,2	1,1	1,1	1,1	1,2†	1,1	1,1	1,1
<i>DQSG5-BII</i>	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,2	1,1	1,2	1,2
<i>DQSG5-BIII</i>	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG6-AI§</i>	1,4	1,3	1,1	3,4	1,3	1,2†	1,1	1,3	1,2†	1,1	1,1	1,1
<i>DQSG7-BI</i>	2,2	1,2	1,2	2,2	2,2	1,2	1,2	1,2	1,1	1,2	2,2	1,1
<i>DQSG7-BII</i>	1,2	2,2	1,2	2,2	1,2	1,2	2,2	1,2	1,2	1,2	1,2	1,2
<i>DQSG7-BIII</i>	1,2	1,1	1,1	1,2	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG8-BI</i>	1,2	1,1	1,2	1,2	1,1	1,1	1,1	1,2	1,1	1,2	1,2	1,2
<i>DQSG8-BII</i>	1,2	1,2	1,1	1,2	2,2	1,1	1,2	1,1	1,1	1,1	1,1	1,1
<i>DQSG8-DI</i>	2,2	1,2	1,2	2,2	2,2	1,1	1,2	2,2	1,2	1,2	1,2	1,2
<i>DQSG8-DII</i>	2,2	1,2	1,2	2,2	2,2	1,1	ND	2,2	1,2	1,2	1,2	1,2
	Family 3						Family 4					
	♂P3	♀P3	G9*	G10	G11	G12	♂P4	♀P4	G14	G15	G16	G17
DNA marker												
<i>DQSG1-AI</i>	1,2	2,2	2,2	2,2	1,2	2,3†	2,2	2,2	2,2	2,2	2,2	2,2
<i>DQSG1-AII</i>	2,2	2,2	2,2	2,2	2,2	2,2	2,2	1,2	1,2	1,2	2,2	2,2
<i>DQSG1-CI</i>	2,4	1,2	1,2	2,4	2,4	2,2	3,3	1,1	1,3	1,3	1,3	1,3
<i>DQSG5-BI</i>	1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,1	1,2	1,2	1,2	1,2
<i>DQSG5-BII</i>	1,1	2,2	1,2	1,2	1,1¶	1,1¶	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG5-BIII</i>	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG6-AI§</i>	1,1	1,2	1,1	1,2	1,1	2,2	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG7-BI</i>	1,2	1,2	2,2	1,1	1,2	1,1	1,2	2,2	1,2	1,2	2,2	2,2
<i>DQSG7-BII</i>	1,1	2,2	1,2	1,2	2,2	1,2	1,2	1,1	1,2	1,2	1,2	1,2
<i>DQSG7-BIII</i>	1,1	1,1	1,2†	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
<i>DQSG8-BI</i>	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,2	1,1	1,2	1,1
<i>DQSG8-BII</i>	1,2	1,1	1,1	1,1	1,1	1,1	1,1	1,2	1,1	1,2	1,1	1,2
<i>DQSG8-DI</i>	1,1	2,2	1,2	2,2	1,1¶	1,1¶	1,1	1,2	1,2	1,1	1,2	1,1
<i>DQSG8-DII</i>	1,2	1,2	1,2	1,1	1,1	2,2	1,2	1,2	1,2	1,2	1,2	2,2

Key: ♂P, ♀P, male and female nest attendants respectively; G, gosling; numbers identify individuals. Numbers in table give alleles for each marker; ND, not determined. * Blue gosling. † One of the putative parents excluded. ‡ Both putative parents excluded. § Allele 4, not previously described, includes a 10.7-kb and a 2.8-kb band. || Putative father excluded. ¶ Putative mother excluded.

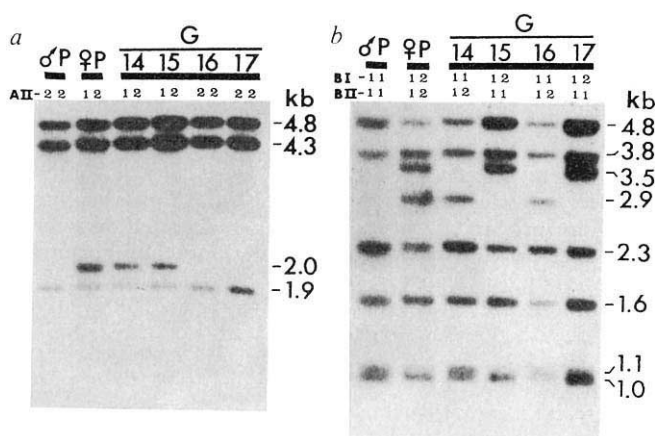


Fig. 1 Mendelian segregation of DNA markers in one snow goose family. Genomic DNA was isolated, digested with either *EcoRI* or *HindIII*, electrophoresed, transferred to GeneScreen Plus membrane and probed as described⁹. The male nest attendant (♂P), female nest attendant (♀P) and goslings (G) are indicated above each track. *a*, *EcoRI* digests of DNA of family 4 probed with *DQSG1*. The alleles AII-1, of 2.0 kb and AII-2 of 1.9 kb are indicated; *b*, *HindIII* digests of DNA of family 4 probed with *DQSG8*. The alleles BI-1 of 2.3 kb and BI-2 of 3.5 kb; BII-1 of 4.8 kb and BII-2 of 2.9 kb are indicated.

Four families with white adult attendants from the breeding colony at La Perouse Bay, Manitoba, Canada (58°24' N, 94°24' W) were studied¹⁰. Two of these families contained blue goslings. Analysis of this plumage colour dimorphism has shown that colour is determined by a single autosomal gene with blue phase dominant to white⁶. This analysis assumed that the presence of blue goslings in nests of white parents (homozygous recessive) was due to IBP and/or EPF. The attendant adults of each family were identified by their behavioural response to the approach of a field-worker. They responded to the presence of the field-worker at the nest by advancing aggressively to about 1.5 m from the nest. As inter-nest distances in the collection areas average 5–8 m, this behavioural response clearly identified the nest attendants. The attendants were shot, their sexes confirmed, and the livers removed and frozen within 1–10 hours. Goslings from the four families were taken from the nest within 24 hours of hatching and were killed immediately before being frozen. DNA was extracted from the livers or the nucleated erythrocytes of the blood⁹. The genomic DNA was digested with one of the four restriction enzymes *HindIII*, *TaqI*, *MspI* and *EcoRI* and the resultant blots probed with a number of the cloned sequences identifying RFLPs.

The genotypes at two DNA marker loci in family 4 are shown in Fig. 1. The male nest attendant was homozygous at locus *DQSG1-AII*, with genotype 2/2, and the female was heterozygous 1/2 (Fig. 1a) (1 and 2 represent different alleles). All four goslings had genotypes compatible with their nest attendants. A similar situation was found at the *DQSG8-BI* and *DQSG8-*

Table 2 Summary of parental exclusions

Family	Gosling no.	Colour phase	DNA marker giving parental exclusion			Conclusion
			♂P	♀P	Either	
1	4	White	<i>DQSG1-CI</i> <i>DQSG8-DI</i> <i>DQSG8-DII</i>	<i>DQSG1-CI</i>	<i>DQSG1-AI</i> <i>DQSG6-AI</i>	IBP
2	5	Blue			<i>DQSG5-BI</i> <i>DQSG6-AI</i> <i>DQSG7-BIII</i>	IBP or EPF
3	9	Blue				IBP or EPF
	10	White	<i>DQSG8-DI</i>	<i>DQSG5-BII</i>		IBP or EPF
	11	White	<i>DQSG7-BII</i>	<i>DQSG8-DI</i>		IBP
	12	White	<i>DQSG6-AI</i>	<i>DQSG8-DI</i> <i>DQSG5-BII</i>	<i>DQSG1-AI</i>	IBP

IBP, intraspecific brood parasitism; EPF, extra-pair fertilization; ♂P, male nest attendant; ♀P, female nest attendant.

BII regions (Fig. 1b). In one case (*DQSG1-CI*, Table 1) each nest attendant was homozygous for a different allele, and all goslings were heterozygous. In total, all 14 DNA markers showed patterns in those 4 goslings compatible with mendelian inheritance of alleles from their nest attendants (parents).

In order to ascertain the origin of the blue gosling in family 2, the alleles at *DQSG6* were analysed (Fig. 2a). The blue gosling (G5) had the genotype 1/2 at the *DQSG6-AI* locus but neither nest attendant carried the 2 allele. At least one of the attendants could not be the biological parent and therefore the blue gosling must have resulted from either IBP or EPF. A similar analysis in family 3 (Fig. 2b) showed that a white gosling (G12) was homozygous, 2/2, for the *DQSG6-AI* marker but the male attendant was homozygous, 1/1. This excludes the male as biological parent but again is consistent with either IBP or EPF.

In order to differentiate between IBP and EPF, the genotypes of all 8 nest attendants and the 17 goslings were determined for a minimum of 13 DNA markers (Table 1). In three families there were a total of six goslings with genotypes that were inconsistent with one or both of the nest attendants (Tables 1 and 2). In family 1, a white gosling (G4) had a genotype indicating both non-paternity and non-maternity, consistent with this gosling resulting from IBP. At *DQSG1-CI*, gosling G4

was homozygous 1/1 whereas the female attendant was 2/2 and the male attendant, 2/3. This marker therefore excludes both attendants from being genetic parents. Similarly at both *DQSG8-DI* and *DQSG8-DII*, gosling G4 was 1/1 but the male attendant was 2/2. At *DQSG1-AI*, gosling G4 was 2/3 but both the male and female attendants were 2/2 (Table 1).

In families 2 and 3, the blue goslings had genotypes inconsistent with one or the other of their nest attendants but could have resulted from either IBP or EPF. Surprisingly, in family 3, three of the four white goslings had genotypes inconsistent with their nest attendants. Of the 17 goslings analysed, six had genotypes which showed that they resulted from either IBP or EPF (Table 2). Four of these were in one family, suggesting that there is variation in susceptibility to IBP or possibly that this represents a pair of nest attendants taking over an abandoned nest containing eggs. At least two of those goslings could not have resulted from EPF. The observation of 6 out of 17 goslings as being illegitimate is higher than that expected based on previous estimates of IBP as being 5–12% in the La Perouse Bay colony¹¹. However, in the year these samples were taken (1985), field observations indicated a particularly high frequency of IBP, particularly in the area where the samples were collected. In addition, two of the broods were chosen non-randomly because of the presence of a blue gosling. This study provides formal proof for the illegitimate origin of these blue goslings in nests of adult white attendants⁶.

Our results clearly demonstrate the power of identifying DNA markers for natural populations where questions of maternity and paternity need to be addressed. Seventeen variable restriction enzyme sites were analysed in this study and when further loci⁹ are examined, an unambiguous differentiation between IBP or EPF of all exceptional goslings will probably be possible.

These studies were supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). T.W.Q. was supported by an NSERC postgraduate scholarship. We thank Dr J. C. Davies for helping with the collection of the samples, and G. D. Schnell, J. N. Thompson, and D. B. Lank who read an earlier version of this manuscript.

Received 7 November 1986; accepted 29 January 1987.

- Hamilton, W. D. *J. theor. Biol.* **7**, 1–52 (1964).
- Mumme, R. L., Koenig, W. D., Zink, R. M. & Marten, J. A. *Auk* **102**, 305–312 (1985).
- Barrowclough, G. F., Johnson, N. K. & Zink, R. M. *Curr. Orn.* **2**, 135–154 (1986).
- Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. *Am. J. hum. Genet.* **32**, 314–331 (1980).
- Jeffreys, A. J., Brookfield, J. F. Y. & Semeonoff, R. *Nature* **317**, 818–819 (1985).
- Cooke, F. & Mirsky, P. J. *Auk* **89**, 863–871 (1972).
- Mineau, P. & Cooke, F. *Behaviour* **70**, 280–291 (1979).
- Bargiello, T. A., Grossfield, J., Steele, R. W. & Cooke, F. *Biochem. Genet.* **15**, 741–763 (1977).
- Quinn, T. W. & White, B. N. *Molec. Biol. Evol.* (in the press).
- Cooke, F., MacInnes, C. D. & Prevett, J. P. *Auk* **92**, 493–510 (1975).
- Finney, G. H. thesis, Queen's Univ., Kingston (1975).

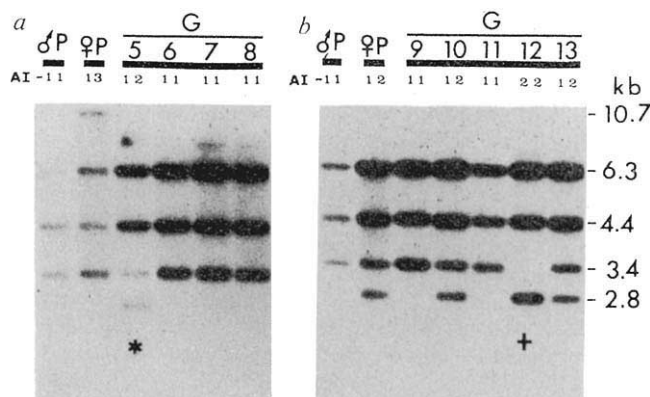


Fig. 2 Parental exclusion in two snow goose families. The Southern blots were prepared as in Fig. 1. a, *EcoRI* digests of DNA of family 2 probed with *DQSG6*. The alleles AI-1, of 6.3 kb, 4.4 kb and 3.4 kb; AI-2 of 6.3 kb, 4.4 kb and 2.8 kb and AI-3 of 10.7 kb and 3.4 kb are indicated; b, *EcoRI* digests of DNA of family 3 probed with *DQSG6*. * Excluded blue gosling (G5) of family 2; +, excluded white gosling (G12) of family 3.

Influence of parental chromosomes on spatial specificity in androgenetic ↔ parthenogenetic chimaeras in the mouse

M. A. H. Surani, S. C. Barton & M. L. Norris

Department of Molecular Embryology,
Institute of Animal Physiology and Genetics Research, Babraham,
Cambridge CB2 4AT, UK

The presence of both parental genomes is essential for development to term in the mouse embryo¹⁻⁵ probably because of germline-specific modifications of homologous chromosomes^{6,7}. Neither androgenetic nor parthenogenetic embryos can by themselves develop to term; any post-implantation embryos they produce have opposite phenotypes, which reflects the presence of complementary information in parental chromosomes^{1,7,8}. The development of androgenetic ↔ parthenogenetic chimaeras is of considerable interest because in this case both parental chromosomes are available even though they are in separate cells. We demonstrate here that in post-implantation chimaeric fetuses, the expression of parental information results in spatial specificity so that parthenogenetic cells are confined to the embryo but the trophoblast consists almost entirely of androgenetic cells. The yolk sac contains both cell types. However, there is incomplete functional complementation because the chimaeras do not reach term. Although failure to reach term may occur partly because of inadequate intermingling and interactions between embryonic cells, it is more likely that further control of mouse development depends on the presence of both sets of chromosomes within the same cells.

Androgenetic and parthenogenetic (equivalent to gynogenetic)⁹ eggs were prepared as described previously^{1,8,10}. The androgenetic eggs contained the paternal genome of strain CFLP, homozygous for the *a* allele of the enzyme glucose phosphate isomerase (GPI), *Gpi-1^a/Gpi-1^a*, and the parthenogenetic eggs contained the maternal (C57BL × CBA)F₁ genome, homozygous for the *b* allele, *Gpi-1^b/Gpi-1^b*. Fertilized eggs homozygous for the two isozyme markers were prepared to serve as controls. All eggs were cultured for approximately 48 h until the 4-cell stage (see Table 1).

The zona pellucida was removed¹¹ and the 4-cell embryos were aggregated¹² to yield androgenetic ↔ parthenogenetic and CFLP × CFLP ↔ F₁ × F₁ chimaeras. The aggregated embryos were cultured for 48 h until the blastocyst stage. All morphologically normal blastocysts were transferred to day-3 pseudopregnant (C57BL × CBA)F₁ females (day 1 = day of vaginal plug) which were prepared by mating with vasectomized F₁ males.

Firstly, the development of androgenetic ↔ parthenogenetic embryos was examined on day 13–15 of gestation. Ninety out of 156 androgenetic eggs reached the 4-cell stage and were aggregated with 4-cell parthenogenetic embryos; 56 of these aggregates formed morphologically normal blastocysts which were transferred into the uterine horns of 21, day-3 recipients. Forty-one implantation sites were detected on day 12–15 of gestation but in all of these the embryos had resorbed with embryonic remains in only 6 of them. By contrast, of the 35 CFLP × CFLP ↔ F₁ × F₁ blastocysts, 16 developed as normal fetuses of which 12 were chimaeric. Hence, although it appeared unlikely that androgenetic ↔ parthenogenetic chimaeras could develop normally to term, some post-implantation development had occurred.

The fate of chimaeras and the spatial distribution of cells was therefore examined on day 10 of gestation when we previously detected the most advanced 6–8-somite androgenetic⁸ and 25-somite parthenogenetic (gynogenetic) embryos with opposite phenotypes^{1,6,8,13,14}. The development of the aggregated

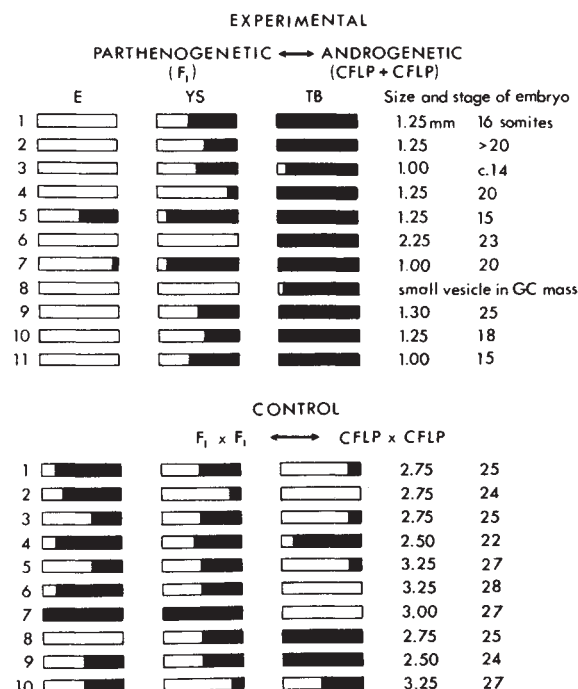


Fig. 1 Proportion of parthenogenetic (open bars) and androgenetic (black bars) cells (with F₁ and CFLP genomes and with GPI types GPI-1B and GPI-1A respectively), in experimental and control embryos. E, embryo; YS, yolk sac; TB, trophoblast. Chimaeric fetuses were obtained either after aggregation of 4-cell androgenetic embryos with 4-cell (Nos 1–7) or 2-cell (Nos 8–11) parthenogenetic embryos.

Methods. The embryos, yolk sacs and trophoblasts were dissected out on day 10 of gestation. The embryos were examined and measured (crown-rump length) and the number of somites recorded. Thereafter, the tissues were stored at -20 °C in lysis buffer until analysed for GPI by cellulose acetate electrophoresis on Helena Titan III electrophoresis plates, essentially as described previously³⁵. The proportion of GPI-1A and GPI-1B in each tissue was estimated by comparison with reference gels, in which different proportions of GPI-1A and GPI-1B were analysed. All non-chimaeric fetuses (see Table 1) were found to be parthenogenetic.

embryos was poor and at best it resembled parthenogenetic fetuses (Table 1, row 1 and Fig. 1, embryos 1–7). Hence, although control embryos were usually 2.5–3.5 mm long with about 25 somites, the experimental fetuses were usually 1.0–1.25 mm with 15–25 somites. The development of the trophoblast was also variable but it tended to be similar to the trophoblast encountered in androgenetic embryos described previously^{6,8}. Further analysis by GPI typing revealed that the embryo was made up almost entirely of parthenogenetic cells and the trophoblast was essentially made up of androgenetic cells (Fig. 1, embryos 1–7). The yolk sac contained both parthenogenetic and androgenetic cells and in some cases the contribution of both cell types to the membrane was almost equal (Figs 1 and 2). By contrast, the distribution of cells in F₁ × F₁ ↔ CFLP × CFLP chimaeric fetuses was essentially random (Fig. 1 and Table 1, row 3).

Previous studies have shown that the early dividing blastomeres make a significantly greater contribution to the embryo whereas the late-dividing blastomeres have a tendency to contribute cells more to the extraembryonic lineage^{12,15}. We therefore aggregated 4-cell androgenetic embryos with 2-cell parthenogenetic embryos in an attempt to force the parthenogenetic cells to contribute to the trophoblast and the androgenetic cells to the embryo (Table 1, row 2). However, parthenogenetic cells were once again confined to the embryo and the yolk sac, whereas the androgenetic cells were found only in the trophoblast and the yolk sac (Fig. 1, embryos 8–11). Therefore there is an inherent tendency of the cells with maternal

chromosomes to contribute to the embryo and those with paternal chromosomes to the trophoblast, but the yolk sac usually contains both cell types. Similar distribution of androgenetic and parthenogenetic cells is also observed in chimaeric fetuses constructed by aggregation with normal embryos (M.A.H.S., S.C.B. and M.L.N., in preparation).

A number of differences between the embryo and some of the extraembryonic tissues have been detected⁷. These include preferential inactivation of the paternal X chromosome¹⁶⁻¹⁹ and the relative undermethylation of some of the CpG DNA sequences²⁰ in the extraembryonic tissues of trophoblastic and primitive endodermal origin. Furthermore, the cellular homologue of the viral oncogene *c-fos* is expressed in day-10 conceptuses predominantly in all the extraembryonic tissue, including the amnion which is of primitive ectodermal origin²¹. Therefore the relationship between the restricted spatial distribution of parthenogenetic and androgenetic cells to the embryo and extraembryonic tissues is of particular interest. Further analysis of the presence of both androgenetic and parthenogenetic cells in the yolk sac, which is a bilayer composed of an inner layer of mesoderm derived from the primitive ectoderm and an outer layer of cells derived from the extraembryonic primitive endoderm cells^{22,23} will clarify the extent to which spatial distribution of cells reflects lineage specificity.

A number of explanations may account for the spatial restriction of androgenetic and parthenogenetic cells in postimplantation fetuses. Slower division rate of androgenetic cells would cause them to remain predominantly in the extraembryonic lineage. However, this can be discounted because no differences were seen even in the 4-cell androgenetic $\leftrightarrow$ 2-cell parthenogenetic chimaeras and because of the relative lack of parthenogenetic cells in the trophoblast under any circumstances. The distribution of cells may therefore be governed by cell surface differences because of different genotypes which could influence cell interactions and allocations. In addition, the observed distribution of cells may be determined by selective proliferation and/or elimination of cells from the embryo or extraembryonic tissues especially between 5.5 and 6.5 days of

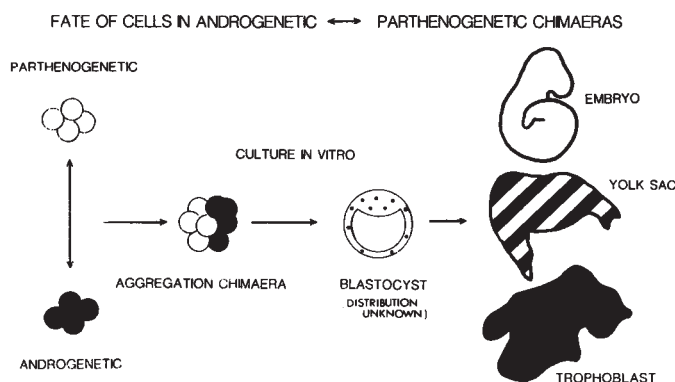


Fig. 2 Preparation of chimaeras and spatial distribution of androgenetic and parthenogenetic cells in chimaeric fetuses.

gestation when regulation in the number of cells and size of the fetus²⁴⁻²⁶ occurs, possibly by alterations in the length of the cell cycle²⁵. Hence, the observed distribution of cells may be dependent on the parental origin of chromosomes through the expression of specific genes implicated in cell growth.

Previous studies have demonstrated the presence of distinct information in the maternal and paternal genomes and their relative importance for the development of the embryo and the extraembryonic tissues, respectively^{1,6,8}. However, despite this distinction there is a lack of functional complementation so that development to term does not occur when the parental genomes are confined to separate cells in androgenetic $\leftrightarrow$ parthenogenetic chimaeras. The presence of both genomes inside some normal cells in which they can interact seems to provide the necessary additional information because fertilized $\leftrightarrow$ parthenogenetic chimaeras develop into adult mice in which parthenogenetic cells contribute to all tissues²⁷⁻²⁹, including germ cells²⁷.

Accumulating evidence from genetic studies which illustrates the importance of the parental origin of specific chromosomal

Table 1 Development of androgenetic $\leftrightarrow$ parthenogenetic chimaeras

Experiment no.	Number			No. of embryos on day 10 of gestation		
	Androgenetic: 4-cell/total prepared	AK ↔ PK* blastocysts/ total aggregates	Transferred to recipients that became pregnant/ total transferred	Implanted	With embryonic derivatives	Chimaeras
1*†	57/173	34/57	26/34	18	12	7
2‡	31/135	29/56	20/29	11	6	4
	Control:					
3	(F ₁ × F ₁ ↔ CFLP × CFLP)	36/38	24/36	18	16	10

* AK, androgenetic; PK, parthenogenetic. Aggregated with 4-cell (*) or 2-cell (‡) parthenogenetic embryos.

For methods, see below and text.

For the preparation of parthenogenetic eggs, cumulus masses containing unfertilized eggs were recovered from (C57BL/6J $\times$ CBA/Ca)F₁ females 16.5–17.00 h after an injection of 5 IU human chorionic gonadotrophin (HCG) in animals which had been injected 48 h previously with 5 IU pregnant mare's serum to induce superovulation. The cumulus cells were dispersed by incubation in 300 IU ml⁻¹ hyaluronidase (ovine testis) in phosphate-buffered medium (PB1) plus 0.4% bovine serum albumin (BSA)³² for 3–5 min at 37 °C. Eggs were activated by culture at room temperature for 4 min in M16 plus 0.4% BSA plus 7% absolute ethanol³³. The eggs were then cultured in M16 plus 0.4% BSA plus 5 μ g ml⁻¹ cytochalasin B for 4 h to suppress second polar body extrusion so producing diploid parthenogenones; after extensive washing in M16 plus 0.4% BSA and a further 4 h culture³⁴ they were carefully checked for diploidy and then cultured until the 2- or 4-cell stage as required. Androgenetic eggs were prepared from F₁ $\times$ CFLP δ fertilized eggs as described previously⁸. The F₁ female pronucleus was first removed from recipient eggs and a second donor CFLP male pronucleus introduced as a karyoplast into the recipient egg by Sendai-virus-assisted fusion¹⁰ to produce CFLP $\times$ CFLP androgenetic eggs which were then cultured for 48 h, when ~20% of the embryos reached the 4-cell stage.

Aggregation chimaeras were prepared as described previously¹². The zona pellucida was removed by exposure of the 4-cell embryos to acid Tyrode's medium, pH 2.0¹¹ and the zona-free androgenetic and parthenogenetic embryos were placed in pairs in separate 50- μ l drops of culture medium. Each pair of embryos was aggregated by physically pushing the two embryos together until the aggregate seemed sufficiently stable. After all the pairs of aggregated embryos had been checked, they were cultured in M16 plus 0.4% BSA at 5% CO₂ in air. All aggregates were rechecked after 2–4 h to ensure that they had not drifted apart. The chimaeric embryos were then left to culture for ~48 h (or 72 h in the case of the asynchronous aggregations) and all those that had developed into morphologically normal blastocysts were transferred to day-3 pseudopregnant recipient females. The recipients were usually autopsied on day 10 of gestation and the implantation sites were carefully dissected to locate the embryos. The description and the typing of the fetuses is given in Fig. 1.

regions for development^{30,31}, together with the embryological data, now provide unequivocal evidence to show that the parental genomes in the mouse are not equivalent. We have previously postulated germline-specific imprinting of chromosomes during gametogenesis which introduces chromosomal determinants of development⁷, which are then propagated after fertilization and used throughout development, at least in the mouse and perhaps in all mammals. Further research may begin to reveal the precise molecular mechanisms of this process which appears to be crucial for mammalian development.

We thank Wolf Reik, Sarah Howlett, Andrew Collick and Nicholas Allen for helpful comments and stimulating discussion.

Received 21 November 1986; accepted 14 January 1987.

1. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Nature* **308**, 548–550 (1984).
2. McGrath, J. & Solter, D. *Cell* **37**, 179–183 (1984).
3. Mann, J. R. & Lovell-Badge, R. H. *Nature* **310**, 66–67 (1984).
4. Anderregg, C. & Markert, C. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6509–6513 (1986).
5. Renard, J. P. & Babinet, Ch. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6883–6886 (1986).
6. Surani, M. A. H., Barton, S. C. & Norris, M. L. *Cell* **45**, 127–136 (1986).
7. Surani, M. A. H., Reik, W., Norris, M. L. & Barton, S. C. *J. Embryol. exp. Morph. suppl.* **97**, 123–136 (1986).

8. Barton, S. C., Surani, M. A. H. & Norris, M. L. *Nature* **311**, 374–376 (1984).
9. Surani, M. A. H. & Barton, S. C. *Science* **222**, 1034–1036 (1983).
10. McGrath, J. & Solter, D. *Science* **220**, 1300–1302 (1983).
11. Nicholson, G. L., Yanagimachi, R. & Yanagimachi, H. *J. Cell Biol.* **66**, 263–274 (1975).
12. Surani, M. A. H. & Barton, S. C. *Dev. Biol.* **102**, 335–342 (1984).
13. Kaufman, M. H., Barton, S. C. & Surani, M. A. H. *Nature* **265**, 53–55 (1977).
14. Kelly, S. J., Mulnard, J. G. & Graham, C. F. *J. Embryol. exp. Morph.* **61**, 331–345 (1978).
15. Takagi, N. & Sasaki, M. *Nature* **256**, 640–641 (1975).
16. West, J. D., Frels, W. I., Chapman, V. M. & Papaioannou, V. E. *Cell* **12**, 873–882 (1977).
17. Harper, M. I., Fosten, M. & Monk, M. *J. Embryol. exp. Morph.* **67**, 127–135 (1982).
18. Lyon, M. F. & Rastan, S. *Differentiation* **26**, 63–67 (1984).
19. Chapman, V., Forrester, L., Sanford, J., Hastie, N. & Rossant, J. *Nature* **307**, 284–286 (1984).
20. Muller, R., Verma, I. M. & Adamson, E. D. *EMBO J.* **2**, 679–685 (1983).
21. Gardner, R. L. & Papaioannou, V. E. in *Early Development of Mammals* (eds Balls, M. & Wild, A. E.) 107–134 (Cambridge University Press, 1975).
22. Gardner, R. L. *J. Embryol. exp. Morph.* **68**, 175–198 (1982).
23. Buehr, M. & McLaren, A. *J. Embryol. exp. Morph.* **31**, 229–234 (1974).
24. Lewis, N. E. & Rossant, J. *J. Embryol. exp. Morph.* **72**, 169–181 (1982).
25. Rands, G. F. *J. Embryol. exp. Morph.* **94**, 139–148 (1986).
26. Stevens, L. C. *Nature* **276**, 266–267 (1978).
27. Stevens, L. C., Varnum, D. S. & Eicher, E. M. *Nature* **269**, 515–517 (1977).
28. Surani, M. A. H., Barton, S. C. & Kaufman, M. H. *Nature* **270**, 601–603 (1977).
29. Searle, A. G. & Beechey, C. V. in *Aneuploidy: Etiology and Mechanisms* (eds Dellarco, V. L., Voytelle, P. E. & Hollaender, A.) 363–376 (Plenum, New York, 1985).
30. Cattanach, B. M. & Kirk, M. *Nature* **315**, 496–498 (1985).
31. Whittingham, D. G. & Wales, R. G. *Aust. J. biol. Sci.* **22**, 1065–1072 (1969).
32. Cuthbertson, K. S. R. *J. exp. Zool.* **226**, 311–314 (1983).
33. Whittingham, D. G. *J. Reprod. Fert. suppl.* **14**, 7–21 (1971).
34. Eicher, E. M. & Washburn, L. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 946–950 (1978).

Depletion of cytosolic free calcium induced by photosynthesis

Anthony J. Miller & Dale Sanders

Biology Department, University of York, Heslington, York YO1 5DD, UK

Cytosolic free calcium is widely believed to behave as a second messenger for a wide variety of metabolic and developmental processes in plant cells^{1–3}. Thus, it is expected that environmental stimuli should provoke a change in the concentration of cytosolic free calcium ($[Ca^{2+}]_c$) which then elicits some further response by the cell. However, although spatial differences in $[Ca^{2+}]_c$ have been observed in plant cells^{4,5}, direct measurement of a temporal change of $[Ca^{2+}]_c$ in response to an environmental stimulus has so far been restricted to the transient rise, largely complete within 2 s, which occurs during an action potential⁶. This paucity of data is largely due to the difficulties associated with measurement of $[Ca^{2+}]_c$ in plant cells⁷. We show here that Ca^{2+} -selective microelectrodes can be used to report photosynthetically-related changes of $[Ca^{2+}]_c$ in the characean alga *Nitellopsis*. We propose that the light-induced depletion of $[Ca^{2+}]_c$ constitutes a fundamental signal which enables the rate of extrachloroplastic metabolism to be geared to photosynthetic processes in the chloroplast. However, light-induced activation of the electrogenic H^+ pump at the plasma membrane is independent of changes in $[Ca^{2+}]_c$.

In order to maximize the tip diameter of the Ca^{2+} -selective electrode, and hence minimize the response time constant, single-barrelled electrodes were employed in conjunction with a KCl-filled potential-measuring electrode. Figure 1 shows that in darkness or dim light, intracellular measurements of $[Ca^{2+}]_c$ fall into two discrete ranges: 10–395 nM, and 0.30–1.56 mM; no values of intracellular $[Ca^{2+}]_c$ were recorded which lay between these spans. The higher range, 'range v', can be interpreted as representing vacuolar Ca^{2+} . The lower of the two, 'range c', is identified with cytoplasmic location of the Ca^{2+} -selective electrode, and measurements in this range also exhibit a clear bimodal distribution (in sub-ranges c_L and c_H). The proportion $c_L/(c_L + c_H)$ (=0.39) is similar to the proportion of observations indicating vacuolar location of the Ca^{2+} -selective electrode, $v/(v + c_L + c_H)$ (=0.42). The simple interpretation, then, is that estimates of cytosolic $[Ca^{2+}]_c$ less than 200 nM result from the vacuolar location of the other, potential-measuring, electrode. This interpretation is consistent with the existence of a small,

inside-positive, trans-tonoplast potential difference⁸. Thus, the true value of $[Ca^{2+}]_c$ in darkness or low-intensity light is in the range 250–400 nM—slightly lower than a previous estimate of $[Ca^{2+}]_c$ (440 nM, assuming a cytosolic free Mg^{2+} concentration of 1 mM), using aequorin in the closely related genus *Nitella*⁶. This small discrepancy might imply that free Mg^{2+} is maintained below 1 mM in the cytosol of characean algae.

Illumination of *Nitellopsis* with white light elicits a decrease in $[Ca^{2+}]_c$ which is reversed by darkness. Figure 2a shows that after illumination by photosynthetically active radiation (PAR) of intensity $65 \mu E m^{-2} s^{-1}$, a stable value of $[Ca^{2+}]_c = 145$ nM is recorded. Darkness then results in a sustained $[Ca^{2+}]_c$ at 390 nM within 1 min. The light-induced decline of $[Ca^{2+}]_c$ was maintained during the measurement times of 5 min and we have not observed any consistent tendency for relaxation of $[Ca^{2+}]_c$ to the dark resting level. As in many other photosynthetic cells, darkness also induces membrane depolarization, which is indicative of inhibition of the plasma membrane electrogenic H^+ pump^{9,10}. However, direct control of electrogenic proton pumping by $[Ca^{2+}]_c$ seems to be ruled out by the different time courses of the two events, particularly after re-illumination. This speculation is confirmed by the experiment of Fig. 2b. The dark-induced rise in $[Ca^{2+}]_c$ from 204 nM to around 350 nM is not reversed by low-intensity light (PAR intensity = $4.5 \mu E m^{-2} s^{-1}$), even though light of this intensity causes membrane repolarization.

The light-intensity dependence of the change in $[Ca^{2+}]_c$ is shown in Fig. 3 (open circles). At the two highest light intensities, $[Ca^{2+}]_c$ declined to a stable value of 55 ± 19 nM (6 observations), which represents a 6–7-fold decline over the steady value in darkness. The half-time for the shift in $[Ca^{2+}]_c$ was variable—sometimes preceding (Fig. 2b) but more often succeeding (Fig. 2a) the light-induced changes in membrane potential. A mean half-time of 35 ± 10 s was recorded for the experiments from which Fig. 3 was derived.

Also shown in Fig. 3 (filled circles) is the light-intensity dependence of carbon fixation, measured in the same conditions on the same cell type. Carbon fixation and the light-induced change in $[Ca^{2+}]_c$ both appear to exhibit a similar, slightly saturating, dependence on light intensity which indicates that the light-induced decrease of $[Ca^{2+}]_c$ is probably related to light absorption by chloroplasts, rather than to some secondary pigment system such as phytochrome¹¹ or a blue-light receptor.

The involvement of photosynthesis in the $[Ca^{2+}]_c$ response is further supported by the results of experiments involving application of the photosystem II inhibitor, 3-(3',4'-dichlorophenyl)-1,1-dimethylurea (DCMU). DCMU completely abolishes the

Fig. 1 Distribution of recorded values in intracellular Ca^{2+} in *Nitellopsis*. Recordings were made in darkness or during illumination with a microscope lamp (PAR intensity $4.5 \mu\text{E m}^{-2} \text{s}^{-1}$). Range 'c' (above) is identified with cytoplasmic location of the Ca^{2+} -selective electrode, and range 'v' (below) with vacuolar location. The sub-range 'c_L' is taken to represent vacuolar location of the potential-measuring electrode (see text), and the *bona fide* recordings of $[\text{Ca}^{2+}]_c$ (both electrode tips located in the cytosol) are therefore given by sub-range 'c_H'.

Methods. *Nitellopsis* sp. was cultured in the laboratory in river mud overlaid with a simple salt solution. Internodal cell segments were prepared²⁵ and kept overnight before use in artificial pond water (APW, composition in mM: K_2SO_4 0.2; NaCl 1.0; CaSO_4 1.0, pH 6.8 with MES-NaOH 5.0). The cell segments exhibited normal rates of cyclosis ($60 \mu\text{m s}^{-1}$) and carbon fixation (Fig. 3). Borosilicate glass micropipettes (tip diameter $1 \mu\text{m}$) were manufactured on a microelectrode puller and filled with 0.1 M KCl (for potential difference recordings) or a Ca^{2+} -selective cocktail containing the ligand ETH 1001 (Fluka). Resistance of the micropipettes when filled with 0.1 M KCl was 5–10 M Ω . Effects of cell turgor on the Ca^{2+} sensor were abolished by incorporating polyvinyl chloride in the cocktail²⁶. Potential-measuring and Ca^{2+} -selective electrodes were positioned in the cell 70–100 μm apart. (This distance compares favourably with the electrical space constant of more than 20 mm for Characean algae²⁷). Measurement of $p\text{Ca}$ was made after subtraction of the signals by a high-input-impedance differential amplifier (World Precision Instruments, Model FD-223). Ca^{2+} electrodes were calibrated in the buffer solutions of Tsien and Rink²⁸, and calibration curves fitted to the data by a non-linear least squares routine²⁹ according to a simplified Nicolsky-Eisenman equation^{25,30} (see Fig. 2, legend). Electrodes frequently exhibited non-Nernstian response from 100–1,000 nM Ca^{2+} , and were not used if the response over this range was $<18 \text{ mV}$. Occasionally, the absolute value of the whole calibration curve shifted by as much as 10 mV after impalement. In those cases, the re-calibration curve was taken as accurate, because comparison of presumed $[\text{Ca}^{2+}]_c$ with values from non-shifting electrodes yielded consistent results only for the second calibration curve. Output from the differential amplifier was sampled at 2 Hz and digitized. On-line display on the monitor was in the format $V=f(t)$ where V is voltage and f some function of time t , or, by applying the fitted calibration curve, $p\text{Ca}_c=f(t)$ or $[\text{Ca}^{2+}]_c=f(t)$. This then linearized output display for electrodes with nonlinear response characteristics over the $p\text{Ca}$ range 6–7. Impalement of voltage and Ca^{2+} microelectrodes into the same cell was performed under the microscope lamp, which would not significantly change $p\text{Ca}$ in Characean algae²⁵. Preparation continuously superfused with APW.

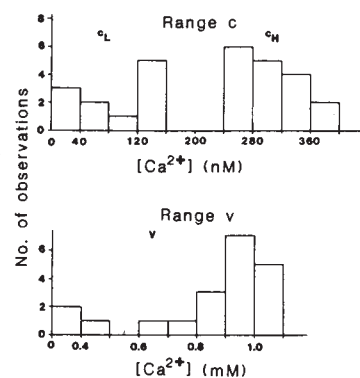
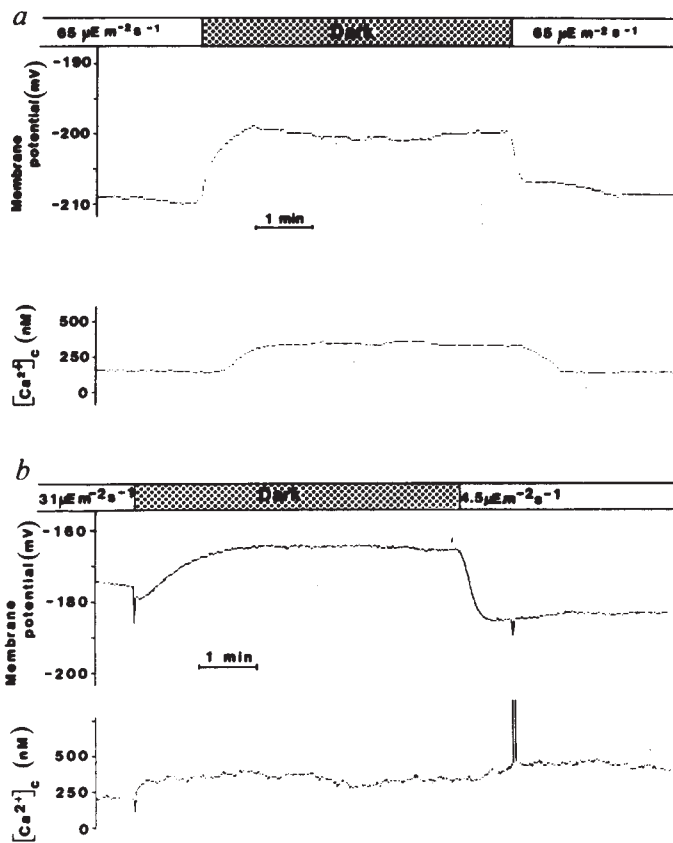


Fig. 2 a, Effect of light/dark transition on membrane potential (upper trace) and $[\text{Ca}^{2+}]_c$ (lower trace). Cell was darkened at time indicated after 6.0 min exposure to white light of PAR intensity $65 \mu\text{E m}^{-2} \text{s}^{-1}$. After 5.5 min darkness, illumination was restored. The apparent changes in $[\text{Ca}^{2+}]_c$ are not significantly rate-limited by the electrode response time: electrode time constant for the $[\text{Ca}^{2+}]_c$ range 10^{-7} – 10^{-6} M is 13 s. Inset, calibration data for Ca^{2+} -selective electrode before (○) and after (●) impalement. The curve, fitted to the post-impalement data, is a simplified Nicolsky-Eisenman equation of the form $E = E_0 + 29.6 \log \{[\text{Ca}^{2+}] + K\}$, in which E is the measured output of the electrode, E_0 is the constant electrode reference potential, and K is a term which subsumes the concentrations of the interfering ions and the selectivity coefficients of the electrode for those ions. Fitted parameter values: $E_0 = 29 \text{ mV}$; $K = 175 \text{ nM}$. b, As a, but for different light intensities. Cell was exposed with PAR of $31 \mu\text{E m}^{-2} \text{s}^{-1}$ for 3.5 min before darkening, and reilluminated with light of lower intensity ($4.5 \mu\text{E m}^{-2} \text{s}^{-1}$) 5.8 min after onset of darkness. The drift present in the $[\text{Ca}^{2+}]_c$ trace in darkness (in this case between extreme values of 280 and 355 nM) was observed occasionally and probably has its origins in low frequency electrode noise. Inset, calibration data as in inset a, with fitted parameter values $E_0 = 94 \text{ mV}$ and $K = 178 \text{ nM}$.



light-induced decrease in $[\text{Ca}^{2+}]_c$ (Fig. 4). (The apparent transient rise in $[\text{Ca}^{2+}]_c$ in Fig. 4 coincides with membrane hyperpolarization and is explained by the relatively large electrical time constant of the Ca^{2+} -selective electrode¹².) Similar data also indicate that application of DCMU in the light causes $[\text{Ca}^{2+}]_c$ to revert to the dark level.

Persistence of the light-induced membrane hyperpolarization in the presence of DCMU is always observed and is explained

if electrogenic H^+ pumping is primarily under the control of cyclic electron transport which proceeds in the presence of DCMU⁹. These results therefore also confirm the conclusion reached above that light-induced activation of electrogenic pumping can be uncoupled from shifts in $[\text{Ca}^{2+}]_c$ and, by implication, is not regulated by $[\text{Ca}^{2+}]_c$ in these conditions. Results from internally-perfused cells of *Nitellopsis*¹³ and *Chara*¹⁴ suggest that any modulation of $[\text{Ca}^{2+}]_c$ below $1 \mu\text{M}$

Fig. 3 Light-intensity dependence of ^{14}C -fixation (●) and light-induced decline in $[\text{Ca}^{2+}]_c$ (○). Fixation of ^{14}C was measured in cell segments in stagnant solutions containing 1 mM NaHCO_3 , but under conditions otherwise identical to those in the electrophysiological experiments. Uptake was for 5 min, and was terminated by placing the cell segment in 0.1 M HCl, which also drove off unfixed CO_2 . The ^{14}C fixation was estimated by scintillation counting after evaporation of the acid. Carbon fixation values are derived after subtraction of a small component of dark fixation amounting to $54 \text{ nmol m}^{-2} \text{ s}^{-1}$. Data points for carbon fixation are the means from 5 cells; bars show s.e.m. For $\Delta[\text{Ca}^{2+}]_c$, points are the means from 2–4 determinations at each light intensity. The line was fitted jointly by method of least squares to the two data sets (appropriately scaled) according to a function of the form $y = ax + bx^2$, with $a = 3.23 \pm 0.48$ and $b = -1.25 \times 10^{-2} \pm 0.68 \times 10^{-2}$ and $x = \text{light intensity}$. Scaling factors: ^{14}C -fixation, $5 \text{ nmol m}^{-2} \text{ s}^{-1}$; $\Delta[\text{Ca}^{2+}]_c$, 2 nM.

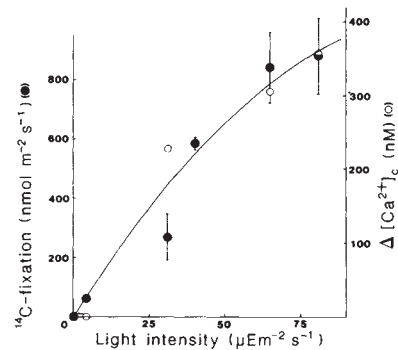
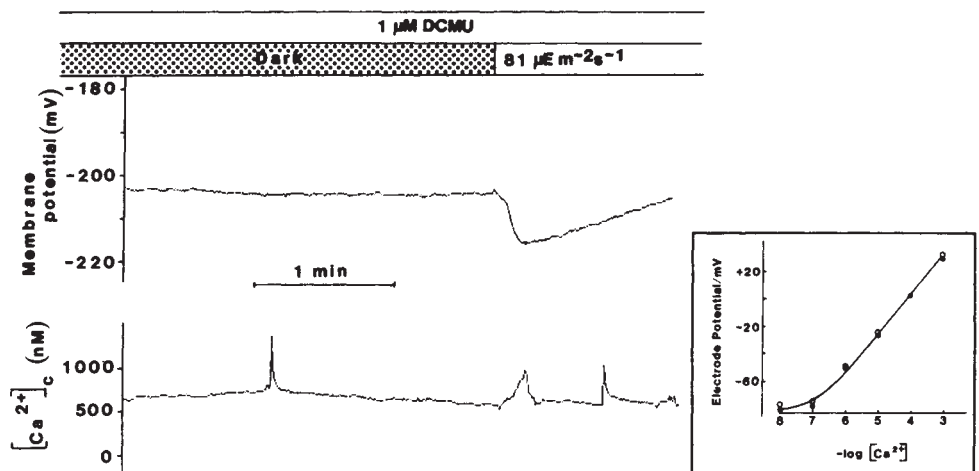


Fig. 4 Effect of dark/light transition on membrane potential and $[\text{Ca}^{2+}]_c$ in the presence of $1 \mu\text{M}$ DCMU. Cell was pretreated with DCMU for 14 min before the start of the trace displayed, and after a further 9 min was impaled with a Ca^{2+} -selective electrode. Application of DCMU caused $[\text{Ca}^{2+}]_c$ to rise (in this case from 280 nM to 610 nM during the period in which the Ca^{2+} -selective electrode was recording intracellularly) in all cells tested. The apparent rise in $[\text{Ca}^{2+}]_c$ after light-on is an artefact caused by disparity between the electrical time constants of the reference and Ca^{2+} -selective electrodes and by the membrane hyperpolarization. Inset, calibration data for Ca^{2+} -selective electrode as for Fig. 2, with fitted parameter values $E_0 = 120 \text{ mV}$ and $K = 154 \text{ nM}$.



has no effect on membrane potential. Because $[\text{Ca}^{2+}]_c$ *in vivo* is maintained in the sub-micromolar range, even in darkness, these perfusion experiments provide support for our conclusion that light-induced shifts in membrane potential are not induced by changes of $[\text{Ca}^{2+}]_c$.

The demonstration that light induces a decline of $[\text{Ca}^{2+}]_c$ in *Nitellopsis* is complemented by reports^{15,16} of light-dependent, DCMU-sensitive Ca^{2+} uptake by isolated intact chloroplasts from spinach. Although Ca^{2+} uptake by chloroplasts is thought to be responsible for light-activation of (stromal) NAD kinase¹⁷, the present results suggest that another consequence of the Ca^{2+} flux should also be considered: the depletion of cytosolic $[\text{Ca}^{2+}]_c$.

One function for light-induced $[\text{Ca}^{2+}]_c$ depletion is the control of sucrose synthesis, which takes place in the cytosol, but which must be closely coordinated with the rate of inorganic carbon reduction in the stroma. Indeed, *Lamprothamnium*, a close relative of *Nitellopsis*, produces considerable quantities of sucrose¹⁸, and for most higher plants sucrose is the dominant form in which reduced carbon is exported from photosynthetically-active cells. Fructose biphosphatase, a key control enzyme in sucrose biosynthesis, is regulated by fructose 2,6-bisphosphate¹⁹ and is also inhibited by micromolar concentrations of Ca^{2+} (ref. 20). It is not yet known whether, in the presence of fructose 2,6-bisphosphate, the enzyme is sensitive to fluctuations of $[\text{Ca}^{2+}]_c$ in the range 50–350 nM for light/dark transitions reported here, nor has the level of Ca^{2+} been strictly controlled by buffering in these *in vitro* investigations. However, it appears probable that the light-induced depletion of $[\text{Ca}^{2+}]_c$ acts in conjunction with fructose 2,6-bisphosphate to accelerate the rate of sucrose synthesis during illumination.

In addition, the dark-induced rise in $[\text{Ca}^{2+}]_c$ can be viewed as a mechanism for activation of respiration in darkness^{21,22}. A

major pathway for oxidation of reducing equivalents in plants is through the NADH oxidase located on the external face of the inner mitochondrial membrane²³ and the respiratory rate of *Helianthus* mitochondria using NADH as substrate doubles as the free Ca^{2+} is raised from 40 to 400 nM (ref. 23). Thus, the dark-induced rise in $[\text{Ca}^{2+}]_c$ may well be involved indirectly in the control of adenine nucleotide levels which are known to be relatively insensitive to light/dark transitions in photosynthetic cells²⁴.

We thank John Thain (University of East Anglia) for supplies of *Nitellopsis*, Ian Jennings for technical assistance, and Richard Firn, Ross Lilley, Tony Moore, Kevin Pyke and Phil Rea for discussions. This work was supported by the AFRC (Grant AG87/26) and the Nuffield Foundation.

Received 22 September 1986; accepted 29 January 1987.

- Hanson, J. B. & Trewavas, A. J. *New Phytol.* **90**, 1–18 (1982).
- Marme, D. & Dieter, P. in *Calcium and Cell Function* Vol. IV (ed. Cheung, W. Y.) 263–311 (Academic, New York, 1983).
- Hepler, P. K. & Wayne, R. O. *A. Rev. Pl. Physiol.* **36**, 397–439 (1985).
- Brownlee, C. & Wood, J. W. *Nature* **320**, 624–626 (1986).
- Keith, C. H., Ratan, R., Maxfield, F. R., Bajer, A. & Shelanski, M. L. *Nature* **316**, 848–850 (1985).
- Williamson, R. E. & Ashley, C. C. *Nature* **296**, 647–651 (1982).
- Clarkson, D. T. in *Molecular and Cellular Aspects of Calcium in Plant Development* (ed. Trewavas, A. J.) 443–447 (Plenum, New York, 1986).
- Spanswick, R. M. & Williams, E. J. *J. exp. Bot.* **15**, 193–200 (1964).
- Spanswick, R. M. *Biochim. biophys. Acta* **332**, 387–398 (1974).
- Hansen, U.-P. *J. Membrane Biol.* **41**, 197–224 (1978).
- Weissenfels, M. H. & Ruppert, H. K. *Planta* **137**, 225–229 (1977).
- Sanders, D. & Slayman, C. L. *J. gen. Physiol.* **80**, 377–402 (1982).
- Mimura, T. & Tazawa, M. *Protoplasma* **118**, 49–55 (1983).
- Sanders, D., Hansen, U.-P. & Gradmann, D. in *Molecular and Cellular Aspects of Calcium in Plant Development* (ed. Trewavas, A. J.) 415–416 (Plenum, New York, 1986).
- Kreimer, G., Melkonian, M., Holtum, J. A. M. & Latzko, E. *Planta* **166**, 515–523 (1985).
- Kreimer, G., Melkonian, M. & Latzko, E. *FEBS Lett.* **180**, 253–258 (1985).
- Muto, S. *FEBS Lett.* **147**, 161–164 (1982).
- Kirst, G. O. & Bisson, M. A. *Planta* **155**, 287–295 (1982).

19. Stitt, M., Hertzog, B. & Heldt, H. W. *Pl. Physiol.* **75**, 548-553 (1984).
20. Foyer, C., Walker, D. & Latzko, E. Z. *Pflanzenphysiol.* **107**, 457-465 (1982).
21. Valles, K. L. M., Proudlove, M. O., Beechey, R. B. & Moore, A. L. *Biochem. Soc. Trans.* **12**, 851-852 (1984).
22. Brooks, A. & Farquhar, G. D. *Planta* **165**, 397-406 (1985).
23. Moore, A. L. & Akerman, K. E. O. *Biochem. biophys. Res. Commun.* **109**, 513-517 (1982).
24. Stitt, M., Lilley, R. M. & Heldt, H. W. *Pl. Physiol.* **70**, 971-977 (1982).
25. Miller, A. J., Parsons, A., Jennings, I. R. & Sanders, D. in *Plant Membrane Transport* (eds Beilby, M. J., Smith, J. R. & Walker, N. A.) (Australian Academy, in the press).
26. Sanders, D. & Miller, A. J. in *Molecular and Cellular Aspects of Calcium in Plant Development* (ed. Trewavas, A. J.) 149-156 (Plenum, New York, 1986).
27. Hogg, J., Williams, E. J. & Johnston, R. J. *J. theor. Biol.* **24**, 317-334 (1969).
28. Tsien, R. Y. & Rink, T. J. *Biochim. biophys. Acta* **599**, 623-628 (1980).
29. Marquardt, D. W. *J. Soc. ind. appl. Math.* **11**, 431-441 (1963).
30. Lee, C. O. in *Ion Selective Microelectrodes and their Use in Excitable Tissues* (eds Syková, E., Huik, P. & Vyklíčí, L.) 47-52 (Plenum, New York, 1981).

The T lymphocyte glycoprotein CD2 binds the cell surface ligand LFA-3

Periasamy Selvaraj, Marian L. Plunkett,
Michael Dustin, Martin E. Sanders*, Stephen Shaw*
& Timothy A. Springer

Laboratory of Membrane Immunochimistry, Dana-Farber Cancer Institute, and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA

* Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

CD2 (known also as T11 (ref. 1), LFA-2 (ref. 2) and the erythrocyte rosette receptor (ref. 3)) is a functionally important T lymphocyte surface glycoprotein of relative molecular mass 50,000 to 58,000⁴ (M_r 50-58 K) which appears early in thymocyte ontogeny and is present on all mature T cells⁵. Monoclonal antibodies to CD2 inhibit cytotoxic T-lymphocyte (CTL)-mediated killing by binding to the T lymphocyte and blocking adhesion to the target cell^{2-4,6}. Such antibodies also inhibit T helper cell responses including antigen-stimulated proliferation, interleukin-2 (IL-2) secretion, and IL-2 receptor expression^{2-4,7-9}. Certain combinations of monoclonal antibodies to CD2 epitopes trigger proliferation of peripheral blood T lymphocytes⁷, cytotoxic effector function¹⁰ and expression of IL-2 receptors by thymocytes, resulting in thymocyte proliferation in the presence of exogenous IL-2 (ref. 11). These findings suggest that CD2 can function in signalling as well as being an adhesion molecule. To understand the role of CD2 in T-cell adhesion and activation, it is essential to define its natural ligand. Our previous observation that purified CD2 inhibits rosetting of T lymphocytes with sheep erythrocytes and can be absorbed by sheep erythrocytes¹² suggested it also might bind with detectable affinity to human cells. We now report that CD2 binds to a cell-surface antigen known as lymphocyte function-associated antigen-3 (LFA-3) with high affinity, and can mediate adhesion of lymphoid cells via interaction with LFA-3.

We established an assay of purified CD2 binding to human lymphoid cells. CD2 was solubilized with the detergent Triton X-100 and purified to near homogeneity from the human T lymphoblast line SKW3 by affinity chromatography as described^{12,13} and labelled with ¹²⁵I (Fig. 1a inset). CD2 was shown to be the major component of this preparation by immunoblotting the unlabelled CD2 preparation with ¹²⁵I-labelled anti-CD2 monoclonal antibody. The sequence of the N-terminus and tryptic peptides of this CD2 preparation are reported elsewhere¹³. CD2 preparations were diluted with buffers containing bovine serum albumin (BSA) (which binds detergent)¹⁴ to a final detergent concentration of 0.001-0.05% depending on the experiment. Under these conditions, at 4 °C, there was no cell lysis. ¹²⁵I-labelled CD2 binding was tested with the B lymphoblastoid cell line JY, which is an excellent target in CTL-mediated killing^{2,4}. Binding of ¹²⁵I-labelled CD2 to JY cells was dependent on the concentration of ¹²⁵I-labelled

CD2 and JY cells. At the highest concentration of JY cells tested, 19% of input ¹²⁵I-labelled CD2 bound (Fig. 1a). Binding was specific, as it was competed out by unlabelled purified CD2 but not by a control preparation of purified lymphocyte function-associated antigen-1 (LFA-1) (Fig. 1b). Binding was half-maximally inhibited at 90 mM CD2, suggesting an association constant of $\sim 10^7$ M⁻¹.

Several recent findings suggested that the ligand for CD2 might be the LFA-3 molecule, a glycoprotein of M_r 55-70K that is broadly distributed on both nonhaematopoietic and haematopoietic cells^{2,4}. Monoclonal antibodies (mAb) against LFA-3 and CD2 inhibit a similar spectrum of antigen-specific helper T and CTL functions⁴. The anti-CD2 antibodies inhibit CTL-mediated killing by binding to the T cells, whereas anti-LFA-3 mAb inhibit by binding to target cells⁴. Both types of antibody block T lymphocyte-target cell adhesion⁶. Studies on antigen-independent conjugation of CTL to B lymphoblastoid target cells have suggested that CD2 and LFA-3 participate in the same adhesion-strengthening functional pathway¹⁵. Antibodies to each antigen partially (about 50%) inhibit conjugate formation. Combinations of saturating concentrations of mAb to LFA-1 and CD2 or LFA-1 and LFA-3 inhibit conjugate formation totally, and thus are additive, whereas the combination of antibodies to CD2 and LFA-3 is no more effective than either antibody alone^{4,15}. CD2 and LFA-3-dependent adhesion can occur in the absence of Mg²⁺ and at 4 °C, whereas LFA-1-dependent adhesion is Mg²⁺ and temperature dependent. Furthermore, thymocyte rosetting with thymic epithelial cells is dependent on CD2 on the thymocyte and LFA-3 on the thymic epithelial cell¹⁶.

Pretreatment of JY cells with the LFA-3 mAb completely inhibited ¹²⁵I-labelled CD2 binding (Table 1). ¹²⁵I-labelled CD2 binding was not inhibited by antibodies against LFA-1, HLA-A,B, or complement receptor type 1 (CR1), but was inhibited by anti-CD2 mAb, confirming the specificity of inhibition and that ¹²⁵I-labelled CD2 molecules had retained the TS2/18 epitope after affinity purification.

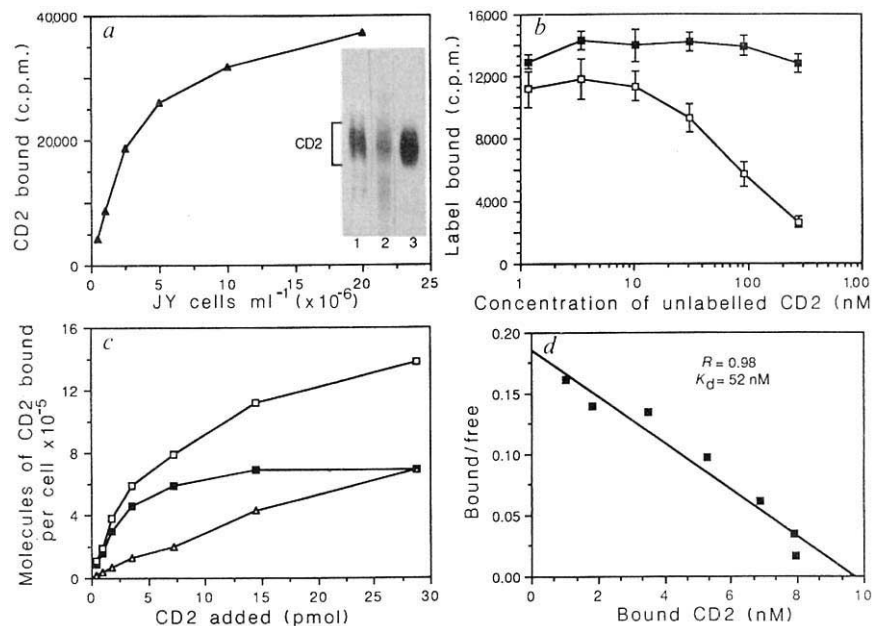
These findings suggest that CD2 and antibodies against LFA-3 compete for binding to cell surface LFA-3 molecules. This was confirmed with the reciprocal approach: inhibition of antibody binding to LFA-3 by CD2 (Fig. 2). JY cells were incubated with purified CD2 or the purified LFA-1 membrane protein as control. The subsequent binding of LFA-3 mAb to cells was determined by immunofluorescence flow cytometry. Purified CD2 inhibited LFA-3 mAb binding by 91% at 420 nM during the preincubation (Fig. 2a,b). Half the maximal inhibition was obtained at 60 nM CD2 (Fig. 2d). Preincubation with CD2 did not affect binding of mAb to intercellular adhesion molecule-1 (ICAM-1), a puta-

Table 1 Inhibition of ¹²⁵I-labelled CD2 binding to JY cells by anti-LFA-3 antibody

Antibody specificity	CD2 bound (mean c.p.m. $\pm$ s.d.)
X63	6,249 $\pm$ 595
Anti-LFA-1	6,900 $\pm$ 493
Anti-HLA-A, B	6,818 $\pm$ 681
Anti-CR1	6,599 $\pm$ 355
Anti-LFA-3	165 $\pm$ 11
Anti-CD2	64 $\pm$ 12

The ¹²⁵I-labelled CD2 was prepared as described (Fig. 1, legend). Specific activity was 3.8×10^8 c.p.m. nmol⁻¹. JY cells (3.3×10^5) were incubated with 50 μ l of antibody (hybridoma culture supernatant) for 45 min at 4 °C. Then 50 μ l of ¹²⁵I-labelled CD2 (diluted to 4,000 c.p.m. μ l⁻¹ with 10% fetal bovine serum (FBS)/RPMI 1640/2 mM HEPES, pH 7.4/3% BSA) was added and the incubation continued for another 2 h at 4 °C. After incubation the cells were washed three times with 10% FBS/RPMI 1640/2 mM HEPES, pH 7.4 and gamma counted. Antibodies used: anti-CD2 (TS2/18), anti-LFA-3(TS2/9), anti-LFA-1(TS1/22)², anti-HLA-A,B(W6/32)²⁷, control (P3X63myeloma IgG₁) and anti-CR1(D44)²⁸.

Fig. 1 125 I-labelled CD2 binding to JY cells. *a*, Binding of 125 I-labelled CD2 to JY cells at various cell concentrations. The input of CD2 was 2×10^5 c.p.m. in 50 μ l, and the binding assay was carried out as described (Table 1 legend). The specific activity of 125 I-labelled CD2 was 3.8×10^8 c.p.m. nmol^{-1} . Inset, purified CD2 subjected to SDS 8% PAGE (polyacrylamide gel electrophoresis) under reducing conditions. Lane 1, polyacrylamide gel after silver staining²⁹; lane 2, autoradiogram of 125 I-labelled CD2; lane 3, autoradiogram after immunoblotting with 125 I-labelled anti-CD2 mAb¹². *b*, Competitive inhibition of 125 I-labelled CD2 binding to JY cells. The specific activity of labelled CD2 was 7×10^7 c.p.m. nmol^{-1} . JY cells (5×10^5 in 100 μ l) were incubated with 125 I-labelled CD2 (10^5 c.p.m.) and unlabelled CD2. The binding assay was as described (Table 1). As a control, purified LFA-1 at the same detergent and buffer concentration as CD2 (but 2.4-fold higher in concentration at each point) was serially diluted in parallel to unlabelled CD2 and mixed with labelled CD2. $\square$, Unlabelled CD2 added; $\blacksquare$, LFA-1 added. *c*, Saturation binding of 125 I-labelled CD2 to JY cells. The specific activity of the 125 I-labelled CD2 was 1.8×10^7 c.p.m. nmol^{-1} . JY cells (5×10^5) were preincubated in 10 μ l of 10% FBS/RPMI 1640 with or without purified anti-LFA-3 antibody (1.3 mg ml^{-1}) for 30 min at 4 °C. Then 50 μ l of varying concentrations of 125 I-labelled CD2 was added and the binding of CD2 was assayed as described above. Nonspecific binding represents binding in presence of anti-LFA-3 mAb. The specific binding was obtained by subtracting nonspecific binding from total binding. Note that at the higher CD2 concentrations used here, the proportion of binding which could not be competed out with antibody to LFA-3 was higher than in other experiments with high specific activity 125 I-labelled CD2 preparations. CD2 molecules bound per cell was calculated taking the M_r of CD2 as 50K. $\blacksquare$, Specific binding; $\square$, total binding; $\triangle$, nonspecific binding. *d*, Scatchard plot. The values obtained as specific binding (Fig. 1c) were used for Scatchard analysis³⁰. The dissociation constant (K_d) was obtained from slope $= -1/K_d$.



Methods. CD2 was purified from the SKW3 T lymphoma cell line by immunoaffinity chromatography as described^{12,13}. The CD2 was eluted from a column of TS2/18 mAb CD2-Sepharose using 0.1 M glycine HCl buffer pH 2.75 containing 0.2 M NaCl and 0.2% Triton X-100. The eluate was immediately neutralized to pH 7.4 with 0.1 vol. of 1 M Tris HCl pH 9.0. Protein was determined on ethanol-precipitated CD2 according to Lowry. Aliquots were labelled with 125 I using 1,3,4,6-tetrachloro-3 α ,6 α -diphenylglycoluril³¹ and extensively dialysed against 10 mM Tris HCl pH 8.0, 0.14 M NaCl, 0.02% NaN_3 . LFA-1 was purified from the same SKW3 cell lysate using a TS1/22 monoclonal antibody²-Sepharose column linked in series to the CD2 antibody-Sepharose column under identical chromatography and elution conditions.

tive ligand for LFA-1¹⁷ (Fig. 2c,d). Preincubation with purified LFA-1 or a matched detergent buffer had no effect on binding of antibodies to either LFA-3 or ICAM-1 (Fig. 2d).

To obtain an accurate measurement of the affinity of soluble CD2 for JY cell surface LFA-3, saturation binding experiments were carried out (Fig. 1c). CD2 showed saturable binding to JY cells. Binding which could not be inhibited by 217 $\mu\text{g ml}^{-1}$ LFA-3 mAb (100-fold the saturating concentration) was taken to be nonspecific and was linearly dependent on the concentration of CD2 added. Scatchard analysis of specific binding shows the association constant $K_a = 1.9 \times 10^7 \text{ M}^{-1}$ (Fig. 1d). At saturation, 680,000 molecules of CD2 were bound per JY cell. The number of CD2 binding sites per JY cell is slightly higher (2- to 3-fold) than an approximate estimate of the number of LFA-3 sites obtained by fluorescence flow cytometry (data not shown). Whether the hydrophobic region of CD2 influences affinity and site number measurements can only be determined when extracellular-domain fragments of CD2 become available.

Functional studies with purified CD2 demonstrate that it mediates adhesion of lymphoid cells (Fig. 3). Although JY cells (LFA-1⁺, ICAM-1⁺, CD2⁺, LFA-3⁺) aggregate at 37 °C by LFA-1-dependent adhesion¹⁸, they do not spontaneously aggregate at 4 °C (Fig. 3a). At 4 °C, the CD2/LFA-3 pathway operates but the LFA-1 pathway does not¹⁵. Addition of purified CD2 (400 nM) resulted in substantial homotypic adhesion of JY cells at 4 °C (Fig. 3b) whereas heat-denatured CD2 was without effect (Fig. 3c). Anti-LFA-3 inhibited adhesion (Fig. 3d). As a specificity control, aggregation induced by the anti-LFA-1 IgM mAb, RDF4 could not be inhibited by anti-LFA-3 mAb (not shown). These results show that purified CD2 can directly mediate cell-cell adhesion by binding to LFA-3. Although we do not

know the physical form of CD2 which mediates cell-cell adhesion in this assay, we believe that a significant amount of CD2 integrates into the JY cell membrane by means of a hydrophobic region at the high concentration we have used for aggregation studies (see the binding which can not be competed by antibody against LFA-3 in Fig. 1c). Sedimentation of CD2 in detergent-free sucrose gradients suggested that the antigen is monomeric (data not shown), and that the aggregation of JY cells was not due to multimeric CD2.

Previous findings that antibodies against CD2 and LFA-3 inhibit T-cell functions and T-cell conjugation to target cells have been interpreted to suggest either that CD2 and LFA-3 are adhesion proteins^{2,4,6,15,16,19,20} or that CD2 and LFA-3 may have a general role in cell function unrelated to adhesion, and that binding of antibody delivers a 'negative signal'^{7,20}. Our results clearly define CD2 as a T-cell adhesion protein. Furthermore, we have found that LFA-3 is a cellular ligand for CD2. Thus both a T-cell adhesion receptor and its ligand have been defined. Thus far, the mechanism of action of other T-cell 'accessory molecules' including LFA-1, CD4 (T4) and CD8 (T8) is unclear. Although functional studies have led to the proposal that they interact with ligands on other cells (ICAM-1, HLA class I, and HLA class II, respectively²¹), in no case has a physical interaction with a ligand, or of a purified molecule with whole cells, been demonstrated. This report is the first demonstration that a purified T-cell surface protein can bind to a cell-surface ligand and mediate cell adhesion.

We are unaware of any previous determinations of the affinity of a cell adhesion molecule for a cell-surface ligand. The affinity of CD2 for LFA-3 is high. The association constant of $1.7 \times 10^7 \text{ M}^{-1}$ is within the range of affinities of antibodies for antigens.

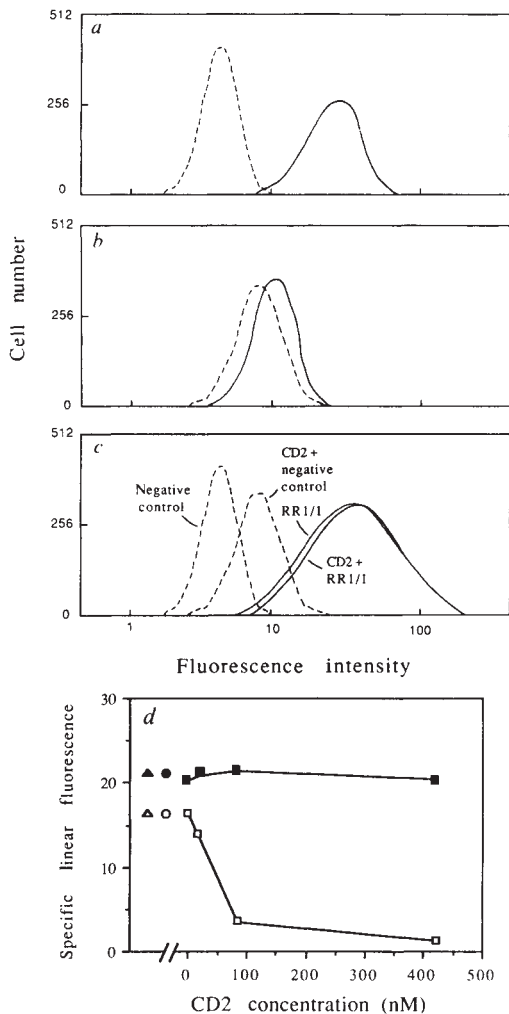


Fig. 2 Inhibition by CD2 of antibody binding to LFA-3. *a*, Anti-LFA-3 mAb (solid line) and nonbinding IgG₁ control antibody (dashed line) staining of JY cells pretreated with control buffer; *b*, as in *a*, but JY cells were pretreated with 420 nM CD2. CD2 caused a concentration-dependent increase in negative control fluorescence which was not given by detergent controls; this may result from nonspecific association of IgG with the hydrophobic portion of bound CD2. *c*, Anti-ICAM-1 mAb RR1/1 (ref. 17; solid line) and nonbinding control antibody (dashed line) staining of JY cells with and without pretreatment with 420 nM CD2. *d*, Inhibition of antibody binding as a function of CD2 concentration. Monoclonal anti-ICAM-1 and LFA-3 antibodies with CD2: ■ and □, respectively; anti-ICAM-1 and anti LFA-3 controls (no detergent): ▲ and △; anti-ICAM-1 and anti-LFA-3 controls with 1,000 nM LFA-1: ● and ○.

Methods. JY cells (10^5) were incubated with purified CD2, purified LFA-1, or control buffer with identical detergent concentration (0.015% or less) in 15% BSA, Hanks balanced salt solution (HBSS) at 4 °C in a volume of 20 μ l. After 60 min, 20 μ l of mAb in HBSS, 15% BSA was added. The lowest concentration of antibody giving optimal staining (2 μ g ml⁻¹ for the monoclonal anti-LFA-3 antibody TS2/9) was used. After 30 min at 4 °C the cells were washed $\times 3$ with HBSS 10% FBS and stained with fluorescein isothiocyanate (FITC)-goat anti-mouse IgG for 30 min at 4 °C. Cells were then analysed on a Coulter Epics V flow cytometer.

Affinities of this order of magnitude may be important in stabilizing adhesion between motile cells and in overcoming repulsion between cells due to net negative surface charge, the glycocalyx and other factors.

We have further found that, as with antigen-independent adhesion of T lymphocytes to target cells, rosetting of activated T lymphocytes with autologous erythrocytes is mediated by the interaction between CD2 and LFA-3 (refs 22, 23). We have also

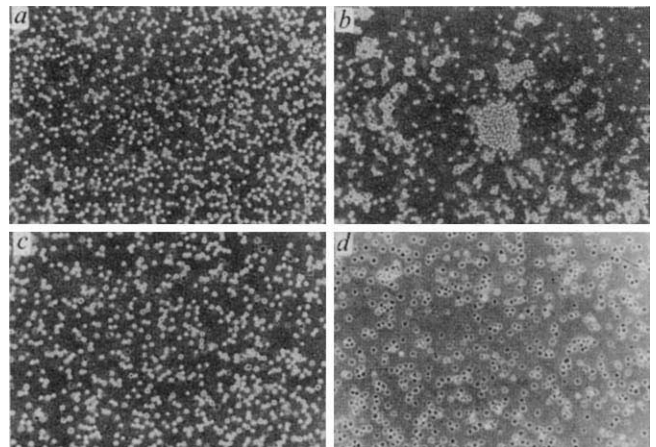


Fig. 3 CD2 and LFA-3-dependent cell adhesion. *a*, JY cells; *b*, JY cells + CD2; *c*, JY cells + heat denatured CD2 (100 °C for 15 min); *d*, JY cells + CD2 + anti-LFA-3 mAb.

Methods. Purified CD2 (84 μ g ml⁻¹) or control buffer (24 μ l), 30% BSA (50 μ l), and 0.25 mg ml⁻¹ antibody (2 μ l) or buffer (2 μ l) were mixed on ice and added to 10^6 JY cells in 24 μ l. The mixture was centrifuged for 2 min at 200g and incubated on ice for 2 h. The cells were gently resuspended and photographed with an inverted microscope at $\times 100$ magnification.

found that purified CD2 mediates aggregation of human erythrocytes, and this is inhibited by antibody to LFA-3 (ref. 22). The interaction between CD2 and the sheep T11TS molecule in rosetting of human T lymphocytes with sheep erythrocytes²⁴. Studies with purified LFA-3 reciprocal to those reported here have confirmed that CD2 is a receptor for LFA-3. Furthermore, we have found that purified LFA-3 reconstituted into planar membranes mediates T lymphocyte adhesion by interaction with CD2²⁵.

Monoclonal antibodies against CD2 can stimulate T-cell proliferation and effector function^{1,10,11}, suggesting that interaction of CD2 with its physiological ligand, LFA-3, may mediate T-cell activation as well as adhesion. Interaction with LFA-3⁺ thymic epithelial cells may be important in driving the proliferation of immature, CD2⁺, antigen-receptor negative thymocytes^{16,21,26}. Further studies are required to determine whether receptor-ligand interaction between CD2 on mature T lymphocytes and LFA-3 stimulates the 'alternative pathway'¹ of T-cell activation or can synergize with the classical antigen receptor pathway.

Supported by NIH grant CA31798 and an American Cancer Society Faculty Award (T.A.S.).

Received 5 November 1986; accepted 29 January 1987.

- Meuer, S. C. *et al.* *Cell* **36**, 897-906 (1984).
- Sanchez-Madrid, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 7489-7493 (1982).
- Van Wauwe, J., Goossens, J., Decock, W., Kung, P. & Goldstein, G. *Immunology* **44**, 865-871 (1981).
- Krensky, A. M. *et al.* *J. Immun.* **131**, 611-616 (1983).
- Haynes, B. F. *Immun. Rev.* **57**, 127-161 (1981).
- Krensky, A. M., Robbins, E., Springer, T. A. & Burakoff, S. J. *J. Immun.* **132**, 2180-2182 (1984).
- Palacios, R. & Martinez-Maza, O. *J. Immun.* **129**, 2479-2485 (1982).
- Reed, J. C., Tadmori, W., Kamoun, M., Koretzky, G. & Nowell, P. C. *J. Immun.* **134**, 1631-1639 (1985).
- Tadmori, W., Reed, J. C., Nowell, P. C. & Kamoun, M. *J. Immun.* **134**, 1709-1716 (1985).
- Siliciano, R. F., Pratt, J. C., Schmidt, R. E., Ritz, J. & Reinherz, E. L. *Nature* **317**, 428-430 (1985).
- Fox, D. A. *et al.* *J. Immun.* **134**, 330-335 (1985).
- Plunkett, M. L. & Springer, T. A. *J. Immun.* **136**, 4181-4187 (1986).
- Selvaraj, P., Plunkett, M. L. & Springer, T. *Biochim. biophys. Acta* (submitted).
- Springer, T. A., Mann, D. L., DeFranco, A. L. & Strominger, J. L. *J. biol. Chem.* **252**, 4682-4693 (1977).
- Shaw, S. *et al.* *Nature* **323**, 262-264 (1986).
- Vollger, L. W., Tuck, D. T., Springer, T. A., Haynes, B. F. & Singer, K. H. *J. Immun.* **138**, 358-363 (1987).
- Rothlein, R., Dustin, M. L., Marlin, S. D. & Springer, T. A. *J. Immun.* **137**, 1270-1274 (1986).
- Rothlein, R. & Springer, T. A. *J. exp. Med.* **163**, 1132-1149 (1986).
- Krensky, A. M. *et al.* *Hybridoma Technology in the Biosciences and Medicine* (ed. Springer, T. A.) 559-573 (Plenum, New York and London, 1985).

20. Barbosa, J. A. *et al.* *J. Immun.* **136**, 3085-3091 (1986).
21. Springer, T. A., Dustin, M. L., Kishimoto, T. K. & Marlin, S. D. *A. Rev. Immun.* **5** (in the press).
22. Plunkett, M. L., Sanders, M. E., Selvaraj, P., Dustin, M. L. & Springer, T. A. *J. exp. Med.* (in the press).
23. Makgoba, M. W., Shaw, S., Gugel, E. A. & Sanders, M. E. *J. Immun.* (submitted).
24. Hünig, T. *J. exp. Med.* **162**, 890-901 (1985).
25. Dustin, M. L., Sanders, M. E., Shaw, S. & Springer, T. A. *J. exp. Med.* (in the press).
26. Haynes, B. F. *Clin. Res.* **34**, 422-431 (1986).
27. Barnstable, C. J. *et al.* *Cell* **14**, 9-20 (1978).
28. Iida, K., Mornaghi, R. & Nussenzweig, V. *J. exp. Med.* **155**, 1427-1438 (1982).
29. Morrissey, J. H. *Analyt. Biochem.* **117**, 307-310 (1981).
30. Scatchard, G. *Ann. N.Y. Acad. Sci.* **51**, 660-672 (1949).
31. Fraker, P. J. & Speck, J. C. *Biochem. biophys. Res. Comm.* **80**, 849-857 (1978).

Compartmentalization of a haematopoietic growth factor (GM-CSF) by glycosaminoglycans in the bone marrow microenvironment

Myrtle Y. Gordon, Graham P. Riley, Suzanne M. Watt & Melyvn F. Greaves

The Leukaemia Research Fund Centre,
The Institute of Cancer Research, Chester Beatty Laboratories,
Fulham Road, London SW3 6JB, UK

Haematopoietic progenitor cells proliferate and mature in semi-solid media when stimulated by exogenous haematopoietic cell growth factors (HCGFs) such as granulocyte-macrophage colony-stimulating factor (GM-CSF)^{1,2}. They also proliferate in association with marrow-derived stromal cells^{3,4} although biologically active amounts of HCGFs cannot be detected in stromal culture supernatants⁵. It is possible that HCGFs are synthesized in small amounts by stromal cells but remain bound to the stromal cells and/or their extracellular matrix (ECM). This interpretation accords with haematopoietic progenitor cell proliferation in close association with stromal layers in long-term cultures⁶. Glycosaminoglycans (GAGs) are found in the ECM produced by stromal cells^{7,8}. They are prime candidates for selectively retaining HCGFs in the stromal layer⁹; they influence embryonic morphogenesis and cyto-differentiation¹⁰ and they may regulate haematopoiesis¹¹⁻¹³. We now report that granulocyte-macrophage colony-stimulating activity can be eluted from cultured stromal layers and that exogenous GM-CSF binds to GAGs from bone marrow stromal ECM. Selective compartmentalization of HCGFs in this manner may be an important function of the marrow microenvironment and may be involved in haematopoietic cell regulation.

Whereas supernatant media collected from bone marrow stromal cell cultures have no detectable HCGF activity, dialysed salt-extracts of the stromal layer itself stimulate colony formation by granulocyte-macrophage colony-forming cells (GM-CFC) in agar cultures (Fig. 1a and b). This result suggested that cryptic HCGF activity is present in marrow stromal cultures and that there are binding sites for growth factors in bone marrow stroma and/or extracellular matrix. We investigated the latter possibility by incubating either medium conditioned by a bladder carcinoma cell line (5637CM)¹⁴ or recombinant GM-CSF (rGM-CSF)¹⁵ with preparations of stromal cultures or purified matrix components and then assaying their residual colony-stimulating activity in agar culture.

Incubation of 5637CM with stromal layers that had been simply dehydrated (D) removed little colony-stimulating activity (Fig. 2a). In contrast, incubation with stromal layers that had been extracted with 2 M NaCl before dehydration (XD) removed 40-60%. Incubation of 5637CM in plates coated with fibronectin alone removed <5% of its colony-stimulating activity but incubation in plates additionally coated with bone marrow (BM)-GAGs removed most of the activity (Fig. 2a).

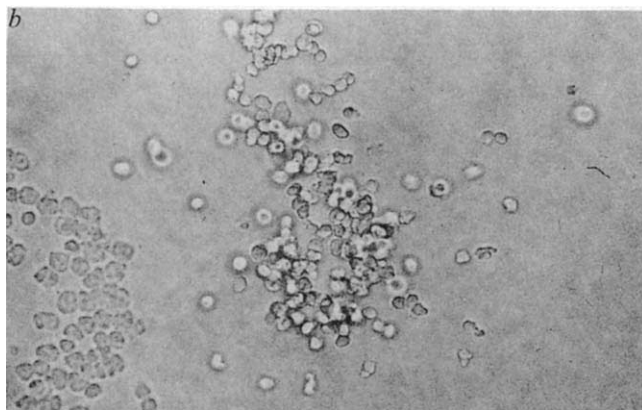
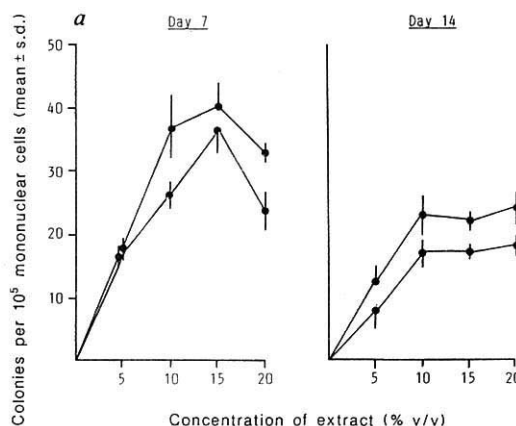


Fig. 1 a, Numbers of day-7 and day-14 GM-CFC stimulated by dialysed salt extracts of intact stromal layers. b, Colonies in 14-day-old agar cultures stimulated by salt extract obtained from intact stromal layers. These colonies contained granulocytes, monocyte-macrophages or mixtures of the two cell types. Colonies shown are ~100 μ m in diameter.

Methods. Confluent stromal layers, grown from 5×10^5 mononuclear cells suspended in 1 ml of α -medium supplemented with 15% fetal calf serum (FCS-GIBCO) and 2×10^{-6} M methylprednisolone (Upjohn) were extracted with 2 M NaCl (30 min at room temperature) and the colony-stimulating activity of the extracts was assayed. Graded concentrations (0-20% v/v) of extract were added to 1 ml 0.3% agar cultures containing 10^5 plastic-adherent cell-depleted mononuclear (separated using Lymphoprep, Nyegaard, Oslo) bone marrow (obtained from informed consenting transplant donors) cells in α -medium (GIBCO) supplemented with 15% FCS. The cultures were incubated at 37 °C in humidified 5% CO₂ in air. Colonies of more than 20 cells were counted after 7 days and after 14 days of incubation. Each line represents results obtained from the marrow of an individual donor.

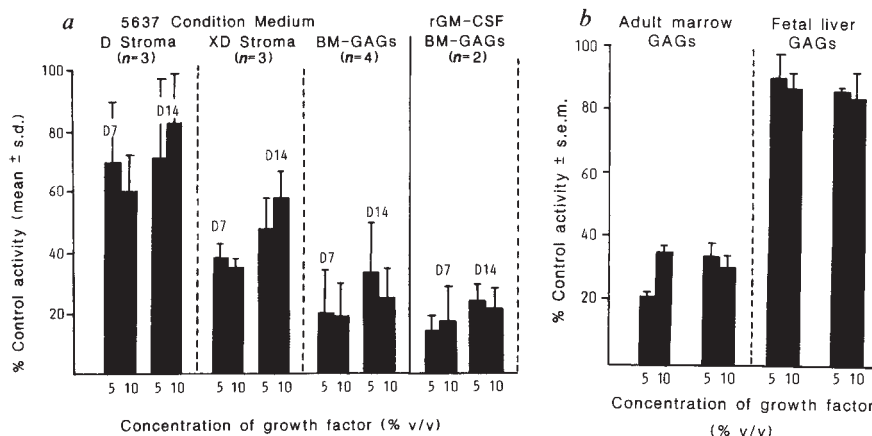
Results of experiments using rGM-CSF and BM-GAGs reproduced those obtained with 5637CM (Fig. 2a).

The reduction in the activity of 5637CM after incubation with BM-GAGs was not caused by the presence of inhibitory material from the coated plates because the addition of α -medium that had been incubated with BM-GAGs to GM-CFC cultures had no significant effect on colony formation stimulated by 5637CM. SDS-polyacrylamide gel electrophoresis (PAGE) analysis of 5637CM before and after incubation with BM-GAGs showed <5% generalized removal of protein despite removal of about 80% of the colony-stimulating activity in <10% of the volume. The presence of a finite number of GM-CSF binding sites on BM-GAGs-coated plates was demonstrated by a titration experiment. Adding increasing amounts of rGM-CSF to BM-GAGs indicated that saturation of binding occurred at 500-600 units per dish (Table 1).

To assess whether the binding of GM-CSF has tissue specificity, we tested the effect of GAGs derived from cultured

Fig. 2 *a*, Reduction in colony-stimulating activity of 5637CM or rGM-CSF post-incubation in plates containing 'D' or 'XD' stroma or BM-GAGs. Results are expressed as % control activity. Vertical bars, standard deviation of the inter-experimental mean; *n*, number of replicate experiments. *b*, Reduction in colony-stimulating activity of 5637CM post-incubation in plates containing GAGs derived from stromal cultures of adult bone marrow (ABM), or fetal liver (FL). The results were obtained from a single experiment; vertical bars, standard error of the ratio²⁷ of the control and experimental values.

Methods. D stromal layers were prepared by dehydrating intact confluent stromal layers in air; XD stromal layers were prepared by extracting the stromal layers with 2 M NaCl for 30 min at room temperature before dehydration. GAGs were prepared from the supernatant media from confluent monolayers of stromal cells using the method of Vannucchi *et al.*²⁷. Petri dishes (35 mm) were coated first with 300 µg of human plasma fibronectin (Sigma) and then with the amount of GAGs extracted from 1 ml of stromal culture supernatant. 5637CM or rGM-CSF (diluted to 1000 units ml⁻¹, that is, equivalent activity to 5637CM) was incubated in the coated dishes overnight or in uncoated dishes as a control, harvested and dialysed against 3 changes of α-medium for 48 hours, sterilized by filtration and then tested for colony-stimulating activity (see legend to Figure 1). Titration of 5637CM and rGM-CSF enabled us to decide that 5% and 10% (v/v) were appropriate concentrations for assaying residual activity in binding studies. The results of two preliminary experiments (data not shown) showed that 5637CM retained more than 95% of its colony-stimulating activity after incubation in plates coated with fibronectin alone.



fetal liver cells (gestational age 18 weeks, at which time the liver is predominantly erythropoietic¹⁶). We found that GAGs from this source were ineffective in removing growth factor from 5637CM (Fig. 2b), although the plates were coated with the same amount of adult marrow and fetal liver GAGs according to scanning densitometry.

Because endogenous colony-stimulating activity could be recovered from stromal layers by salt extraction we attempted to recover similarly exogenous GM-CSF bound to BM-GAGs. Dialysed salt extracts of BM-GAGs that had previously been incubated with 5637CM possessed a level of granulocyte-macrophage colony-stimulating activity that was, on average, 66% of the activity removed by binding (Fig. 3). Salt-extracts of BM-GAGs that had been incubated with α-medium or α-medium + 15% fetal calf serum (FCS) did not stimulate any colony-forming cells. We next tested whether exogenous GM-CSF bound to BM stroma was active *in situ*. Dehydrated (D) stroma had endogenous HCGF activity which was approximately halved by salt extraction (XD stroma) (Table 2). The latter procedure releases functional HCGF as shown earlier (Fig. 1). Binding of 5637CM to XD stroma reconstituted their capacity to stimulate GM-CFC to that of D stroma (Table 2).

Taken together, these data indicate that GM-CSF binds to adult BM stroma and to stroma-derived GAGs. Moreover,

bound growth factor appears to be functional *in situ* and can be eluted from these matrix proteins with retention of biological activity. The physiological relevance of these observations is reinforced by our preliminary finding that dehydrated stromal layers can directly stimulate colony formation and that endogenous growth factor can be removed by salt extraction. The relationship of this putative HCGF to GM-CSF or other well-characterized factors remains to be investigated. It is noteworthy, however, that the salt-extracted activity is not neutralized by antibody to rGM-CSF (S.M.W., unpublished observations).

There is evidence from other studies that specific growth factors are accumulated by the appropriate tissues *in vivo*. A transferrin-like growth factor, which stimulates muscle cells *in vitro*, is selectively accumulated by, and can be eluted from, chicken skeletal muscle during embryonic development¹⁷ and binds to heparin¹⁸. More recently, a heparin-binding endothelial growth factor has been extracted from endothelial cell cultures¹⁹. These studies support the idea that GAGs are involved in concentrating the appropriate amount of growth factors in the relevant target tissues. However, we are not aware of any other studies documenting either the accumulation of exogenous growth factors by matrix glycoproteins or the binding of growth factors to substances produced by one class of cells (marrow

Table 1 Saturation of BM-GAGs with rGM-CSF

Units ml ⁻¹ rGM-CSF incubated with BM-GAGs	Relative activity (%) after incubation		Units rGM-CSF absorbed	
	Day 7 colonies	Day 14 colonies	Day 7 colonies	Day 14 colonies
625	4 ± 1	15 ± 1	600	529
1,250	47 ± 5	59 ± 3	588	518
2,500	74 ± 5	75 ± 2	650	633
5,000	89 ± 3	88 ± 3	550	590

Increasing concentrations (first column) of rGM-CSF were incubated in plates coated with BM-GAGs, as described in the text. Post-incubation, the rGM-CSF samples were diluted to the same concentration for bio-assay and their activities compared with that of a control sample, incubated without BM-GAGs, for calculation of the residual activity shown in the second and third column. The units of rGM-CSF absorbed at each concentration were calculated from the residual activity.

Table 2 Colony stimulating activity of growth factor bound to matrix

Substrate	Added growth factor	Colonies per adherent fraction of mononuclear cells 5 × 10 ⁵ (mean ± s.d.)	
		Experiment 1	Experiment 2
D stroma	α-medium	37.8 ± 4.3	43.0 ± 7.0
XD stroma	α-medium	20.8 ± 7.2	18.8 ± 8.0
XD stroma	Bound activity from 5637CM	45.8 ± 7.0	40.3 ± 3.9

Stromal layers were dehydrated or salt extracted and dehydrated as described in the legend to Fig. 2a. Stromal layers (D and XD) were rehydrated in α-medium or 5637CM as a source of GM-CSF, and the excess fluid removed, before addition of bone marrow cells. Bone marrow mononuclear cells were added to rehydrated D and XD stromal layers. The stromal layers were washed after 2 h and overlaid with 1 ml 0.3% agar in α-medium + 15% FCS. Colonies of early granulocytic and monocytoid cells produced by the adherent cells were scored after 7 days incubation at 37 °C in 5% CO₂ in air.

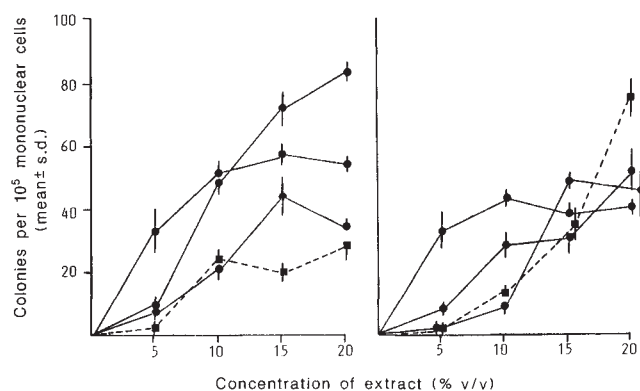


Fig. 3 Colony-stimulating activity extracted from BM-GAGs after incubation with (●) 5637CM or (■) rGM-CSF. Salt extracts (2 M NaCl) were made from plates coated with BM-GAGs that had been incubated with 5637CM or rGM-CSF. The extracts were dialysed and their colony-stimulating activity was assayed in agar cultures.

stromal cells) but acting on another class of cells (haematopoietic cells).

Detailed biochemical analysis of human and mouse bone marrow stroma derived glycosaminoglycans indicates that the major constituents are heparan sulphate, hyaluronic acid and chondroitin sulphate²⁰⁻²². Preliminary analysis suggests that our human BM-derived GAGs which absorb HCGF consist predominantly of hyaluronic acid with a trace of chondroitin sulphate (G.P.R., unpublished observations). More detailed biochemical analysis will be required to determine whether hyaluronic acid or some minor component is responsible for binding growth factors. So far, we have been unable to reproduce the effects of BM-GAGs using chondroitin-4- or 6-sulphates from whale or shark cartilage or hyaluronic acid from human umbilical cord. This is perhaps not surprising in view of the origin of these purified glycoproteins and their known heterogeneity^{23,24}. The diversity of GAGs may indeed equip them for specific recognition events.

Our observations may help to explain some apparent discrepancies in the proposed role of haematopoietic growth factors. For example, GM-CSF is believed to be largely responsible for the positive regulation of GM-CFC *in vivo*¹, but there is little evidence that the endogenous levels of GM-CSF that can be measured in serum and in tissues bear any close relationship to levels of granulopoietic and monocytopenic activity in the bone marrow (reviewed in ref. 2). In addition, granulopoiesis is more than adequately supported by the conditions provided in long-term bone marrow cultures³, yet interleukin 3 (IL-3), GM-CSF and G-CSF are not detectable (ref. 5 and T. M. Dexter, personal communication). One plausible explanation is that the amounts of colony stimulating factors (including GM-CSF) present in a variety of tissues¹ provide for a rapid local stimulation of granulocytes and monocytes, usually in response to infection. In contrast, direct cell-cell transfer of HCGFs may occur between stromal and haematopoietic cells in long-term cultures or high local extracellular concentrations of HCGFs may be produced in some way²⁵. The present findings support the latter explanation and provide a possible mechanism that could be responsible for local accumulation of HCGFs in the marrow *in vivo* as well as in long-term cultures.

In conclusion, we have demonstrated the ability of glycosaminoglycans from the marrow microenvironment to accumulate HCGFs. We suggest that this capacity mirrors an important function of the extracellular matrix in the microenvironment, both *in vivo* and *in vitro*, namely, to compartmentalize growth factors synthesized locally by stromal cells or produced elsewhere and to 'present' them in a biologically active form to haematopoietic progenitor cells.

The work was supported by the Leukaemia Research Fund of Great Britain. We thank the Bone Marrow Transplantation Teams of the Royal Marsden and Hammersmith Hospitals for providing the bone marrow samples, Dr J. Delamarter for the rGM-CSF and Dr G. Morstyn for the anti-rGM-CSF monoclonal antibody. Miss Geraldine Parkins prepared the manuscript.

Received 12 December 1986; accepted 20 January 1987.

1. Metcalf, D. *Blood* **67**, 257-267 (1986).
2. Gordon, M. Y., & Barrett, A. J. *Bone Marrow Disorders: the Biological Basis of Clinical Problems* (Blackwell, Oxford, 1984).
3. Dexter, T. M., Allen, T. D. & Lajtha, L. G. *J. Cell Physiol.* **91**, 335-344 (1977).
4. Gordon, M. Y., Hibbin, J. H., Kearney, L. U., Gordon-Smith, E. C. & Goldman, J. M. *Brit. J. Haemat.* **60**, 129-1356 (1985).
5. Castro-Malaspina, H. *et al. Blood* **56**, 289-301 (1980).
6. Coulombel, L., Eaves, A. C. & Eaves, C. J. *Blood* **62**, 291-297 (1983).
7. Reincke, U., Hsieh, P., Mauch, P., Hellman, S. & Chen, L. B. *J. Histochem. Cytochem.* **30**, 235-239 (1982).
8. Zuckerman, K. S. & Rhodes, R. K. in *Hematopoietic Stem Cell Physiology* (eds Cronkite, E. P., Dainiak, N., McCaffrey, R. P., Palek, J. & Quesenberry, P. J.) 257-266 (Liss, New York, 1985).
9. Muir, H. *Trans. Biochem. Soc.* **11**, 613-622 (1983).
10. Toole, B. P. in *Neuronal Reorganization* (ed. Barondes, S. H.) 275-329 (Plenum, New York, 1976).
11. McCuskey, R. S. & Meinke, H. A. *Am. J. Anat.* **137**, 187-193 (1973).
12. Schrock, L. M., Judd, J. T., Meinke, H. A. & McCuskey, R. S. *Proc. Soc. exp. Biol. Med.* **144**, 593-599 (1973).
13. Pleomacher, R. E., Van't Hull, E. & Van Soest, P. L. *Expl. Hemat.* **6**, 311-318 (1978).
14. Myers, C. D., Katz, F. E., Joshi, G. & Millar, J. L. *Blood* **64**, 152-156 (1984).
15. Burgess, A. W. *et al. Blood* (in the press).
16. Gilmour, J. R. *J. Path.* **52**, 25-36 (1941).
17. Vannucchi, S., Fibbi, G., Cappelleri, R., Del Rosso, M. & Chiarugi, V. *J. cell Physiol.* **111**, 149-154 (1982).
18. Matsuda, R., Spector, D., Micon-Eastwood, J. & Strohmman, R. L. *Dev. Biol.* **103**, 276-284 (1984).
19. Kardami, E., Spector, D. & Strohmman, R. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8044-8047 (1985).
20. Vlodavsky, I., Fridman, R., Sasse, J., Folkman, J. & Klagsbrun, M. *J. cell. Biochem. suppl.* **10**, C 190a (1986).
21. Wight, T. N., Kinsella, M. G., Keating, A. & Singer, J. W. *Blood* **67**, 1333-1343 (1986).
22. DelRosso, M. *et al. Biochem. biophys. Acta* **676**, 129-136 (1981).
23. Gallagher, J. T., Spooner, E. & Dexter, T. M. *J. cell Sci.* **63**, 155-171 (1983).
24. Dietrich, C. P. & Braz, J. *Med. Biol. Res.* **17**, 5-15 (1984).
25. Gallagher, J. T., Lyon, M. & Steward, W. P. *Biochem. J.* **236**, 313-325 (1986).
26. Dexter, T. M., Spooner, E., Vargon, J., Allen, T. D. & Lanotte, M. in *Haematopoietic Stem Cells* (eds Killman, Sv. Aa., Cronkite, E. P. & Muller-Berat, C. N.) 303-322 (Munksgaard, Copenhagen, 1983).
27. Blackett, N. M. *Cell Tissue Kinet.* **7**, 407-412 (1974).

Somatic hypermutation of an immunoglobulin transgene in κ transgenic mice

R. L. O'Brien*†, R. L. Brinster† & U. Storb*‡§

* Department of Microbiology and Immunology, SC-42, University of Washington, Seattle, Washington 98195, USA
† School of Veterinary Medicine, University of Pennsylvania, Philadelphia 19104, USA

Initial studies of somatically acquired mutations in immunoglobulin V regions from hybridomas and myelomas that are not derived from joining aberrations, suggested a controlled and specific hypermutation process¹⁻⁹, because spontaneous mutation rates observed for other genes are extremely low¹⁰. Some evidence for the idea that mutations are introduced during V-gene rearrangement came from the clustering of mutations at the joining sites^{6,8}, from the absence of mutations in unrearranged V genes^{4,5,11} and from the low level of mutations in only partially (D-J) rearranged nonproductive heavy-chain alleles¹². Another model in which mutations accumulate with each cell division, rather

‡ Present addresses: National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, Colorado 80206, USA (R.L.O.B.); Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th Street, Chicago, Illinois 60637, USA (U.S.).

§ To whom correspondence should be addressed.

than being introduced all at once⁸, was supported by the finding that immunoglobulin genes of hybridomas derived from a single mouse frequently had several mutations in common, and so might be derived from the same precursor cell whose daughters then accumulated additional mutations¹³⁻¹⁷. But the common mutations in some cases could be due to as yet unidentified related germline genes, or could represent the effect of antigen selection for certain amino acids. To try to detect hypermutation in the absence of *V*-gene rearrangement, we isolated B lymphocytes with endogenous heavy-chain gene mutations from transgenic mice carrying pre-rearranged κ -transgenes. We found that these κ -transgenes were also somatically mutated. This and other observations indicated that: (1) ongoing rearrangement is not required for mutation; (2) there are signals for hypermutation in the transgenes; (3) the mutations are found only in the variable region, so the constant region may not be a target; (4) different transgene insertion sites are compatible with hypermutations and (5) more than one transgene is expressed in the same cell.

We used antigenic selection to increase the likelihood of isolating cells that had undergone somatic mutation in transgenic mice carrying the rearranged κ -gene isolated from the myeloma MOPC-167¹⁸. The light-chain V region encoded by this gene (*V_κ*167) is known to be frequently used in the antibody response to the hapten phosphorylcholine (PC). Moreover, only one germline *V* gene is present in the *V_κ*167 family^{4,19}. The possibility of confusing a rearranged endogenous *V_κ*167 gene with the transgene is avoided by marker mutations that differ from the germline sequence in the transgene. Moreover, most anti-PC antibodies use a single *V_H* gene, *T₁₅V_H*⁶, making an endogenous gene of known sequence available in the same anti-PC hybridomas.

Mice transgenic for the rearranged κ -gene of MOPC-167 were hyperimmunized with PC-keyhole limpet haemocyanin (KLH). The mice were used to prepare hybridomas, which were then screened for reactivity to the same PC hapten coupled to bovine serum albumin (BSA). The *T₁₅V_H* endogenous gene⁶ served as an internal cellular control for somatic mutation in a gene that had undergone rearrangement, as opposed to the pre-rearranged transgene. Both endogenous *V_H* genes and the κ transgenes could then be sequenced to test for mutations.

Seven anti-PC hybridomas, from two unrelated MOPC-167- κ transgenic mice, were selected for in-depth analysis (see Table 1). Southern blots confirmed that the κ transgene is present in all the hybridomas (not shown). When probed with a *T₁₅V_H* probe, all hybridomas have the 1.1-kilobase (kb) rearranged *Bam*HI band characteristic of rearrangement of this endogenous *V_H* gene (not shown). RNA analysis indicated that both the κ -transgene and the *T₁₅V_H* gene were transcribed in all these hybridomas. An enzyme-linked immunosorbent assay (ELISA) demonstrated that four of the hybridomas produced immunoglobulin G (IgG); two of these (3A1 and 2F2) later appeared to be identical (see below), so that only three IgG hybridomas are shown in the figures. Western blots of SDS-polyacrylamide reducing gels corroborated the heavy-chain class determinations, and also gave presumptive identification of a transgene light-chain product, because the MOPC-167 κ -chain migrates somewhat slower in a 10% gel than do many other κ -chains (not shown). Thus, these hybridomas are suitable test subjects for somatic mutation by virtue of their retention and expression of the κ -transgene, rearrangement and transcription of an endogenous *V_H* gene of known sequence (*T₁₅V_H*), and the use by some of a heavy-chain class frequently associated with somatic mutation (IgG).

Poly(A)⁺ mRNA preparations from the seven hybridomas described above were used to synthesize specifically-primed complementary DNA for *V_H* sequence analysis (Fig. 1). The three IgG hybridomas each differed in sequence from the four identical IgMs; thus all IgG-producing hybridomas appear to bear somatic mutations. (An eighth hybridoma, 2F5, had a *V_H* sequence identical to 3A1, and thus 2F5 and 3A1 are probably

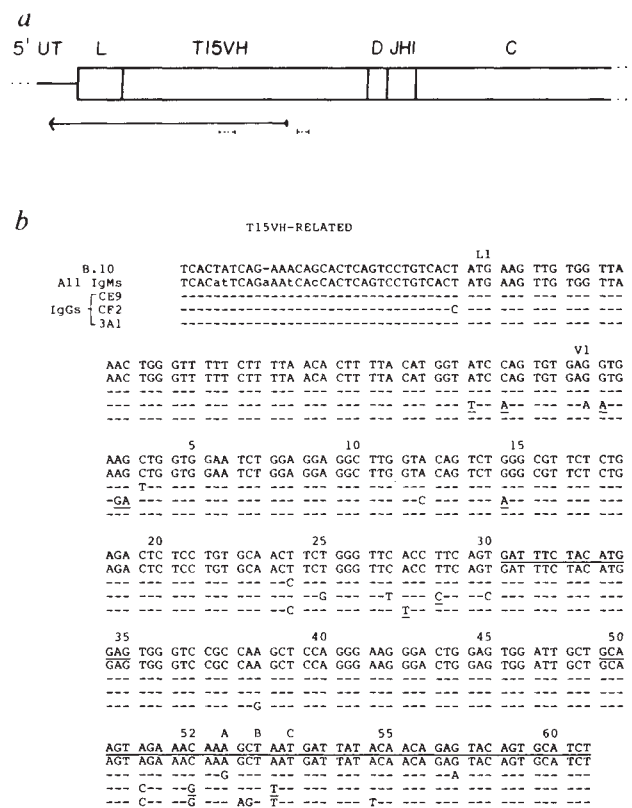


Fig. 1 Heavy-chain mRNAs. *a*, Heavy-chain mRNA in hybridomas: 5', UT 5' untranslated region; L, leader peptide; *T₁₅V_H*, D, JH1 and C genes as marked. Area sequenced is shown by an arrow. Sequencing was by chemical degradation³³ of end-labelled cDNAs, prepared by primer extension of mRNA (F. Burns, personal communication). Primers used are shown as dashed bars. *b*, Sequences of mRNA from eight hybridomas. The sequence of the *T₁₅V_H* gene of C57BL/10 mice (B.10) is shown on top line for comparison. Four IgM-hybridomas differed from this sequence at five places, shown as lower case letters, in the 5' untranslated region. IgGs matched IgMs at all points except where indicated; a dash indicates identity. Changes resulting in a different amino acid are underlined. Complementarity-determining regions are underscored, and amino acids numbered²⁴. Hybridoma 2F5 had an identical sequence to that shown for 3A1.

subclones of the same original cell.) The sequences shown are compared to the germline sequence for the C57BL/10 *T₁₅V_H*-like gene (B.10)²⁰; sequences from both the IgM and IgG hybridomas differ slightly from B10 in the 5' untranslated region. The genes whose transcripts are shown could originate from a heavy-chain allele in either of two mouse strains making up the genetic background of the transgenic mice used: C57BL/6 or SJL. The C57BL/6 strain, like BALB/c mice, appears to use only one *T₁₅V_H*-like gene in the anti-PC response²⁰; we assume this is also true of SJL mice because they are of the same IgH-V allotype as C57BL/6²¹. Whether their *V_H* alleles are identical, however, is not known, and some of the common mutations in Fig. 1*b* could be explained by allelic differences. One is present in CE9 and 3A1 in amino-acid position 24; another in CF2 and 3A1 in amino-acid positions 51, 52 and 52C. One, but not both, could represent an allele different from the one expressed by the IgM-producing hybridomas, assuming that the sequence from the IgM hybridomas represents one of two possible alleles. Neither of the two cases can be due to a prior clonal relationship between the cells, because the cells were derived from different mice. We speculate that one or both cases may indicate a possible gene conversion-like recombination mechanism for some somatic mutations. In any event, we conclude that the three IgG

Table 1 Summary of hybridoma characteristics

Cell	Mouse of origin	Transgene copy no.	V _κ 167 RNA	T ₁₅ V _H RNA	Anti-PC-BSA ELISA	Anti-IgM ELISA	Anti-IgG ELISA
CE9	233-8-5	13	+	+	+	—	+
CF2	233-8-5	13	+	+	+	—	+
2F5	194-2BM.3	3	+	+	+	—	+
3A1	194-2BM.3	3	+	+	+	—	+
1C3	194-2BM.3	3	+	+	+	+	—
3G10	194-2BM.3	3	+	+	+	+	—
3H11	194-2BM.3	3	+	+	+	+	—
4E10	194-2BM.3	3	+	+	+	+	—

Mice transgenic for the κ chain of MOPC-167¹⁸ were prescreened, using an ELISA assay, for low serum titres of anti-PC antibody, because we have often observed high spontaneous anti-PC levels in these mice (R.L.O'B. and J. T. Manz, unpublished data). Two low-serum-titre mice were injected (intraperitoneally) with 100 μ g of PC coupled to keyhole limpet haemocyanin (PC-KLH) in complete Freund's adjuvant, boosted about 3 weeks later with 200 μ g PC-KLH in incomplete Freund's adjuvant, and a third time with 100–200 μ g PC-KLH in saline 14–17 d later. A final injection of PC-KLH in saline was administered one day before fusion with Ag8.653 cells (ATCC)²⁹. Hybridomas were prepared and selected as described³⁰ and cloned by limiting dilution one to two times. Two different mice were used: 233-8-5, which contains about thirteen copies of the κ gene with a small tag (~150 bases) of pUC18 sequence; and 194-2BM.3, which contains 2–3 copies of the κ gene with the entire sequence of pUC18¹⁸. Hybridomas CE9 and CF2 were derived from mouse 233-8-5; all others described here are from 194-2BM.3. Both mice have mixed C57BL/6 and SJL genetic backgrounds. Two anti-PC hybridomas were isolated from 233-8-5, and both produced IgG. Twenty-four anti-PC hybridomas were isolated from 194-2BM.3; two produced IgG, and the rest IgM. Anti-phosphorylcholine (PC) hybridomas were detected using PC-BSA coated microtitre dishes³¹, by adding 50 μ l of hybridoma supernatant per well and detecting bound hybridoma antibody with a peroxidase-conjugated goat anti-mouse Ig antibody preparation (Kirkegaard and Perry Labs). IgG and IgM subclasses were determined a similar way, by using goat anti-mouse IgG or IgM antisera to first coat the wells (Cappel). Three to four ml of each hybridoma culture, approaching saturation, was used for RNA dot analysis³²; 1/10th of the sample was used for each dot. Duplicate dots were hybridized to nick-translated probes of the variable region of the MOPC-167 transgene (V_κ 167) and a heavy-chain variable region of the T₁₅ V_H type (V_H 167), which is most often used in the anti-PC response⁶.

hybridomas bear somatic mutations in the rearranged endogenous V_H gene, but the IgM hybridomas show no evidence for mutation. We then tested the κ -transgenes in the IgG-producing hybridomas for immunoglobulin gene somatic hypermutation, knowing that this event had occurred in endogenous genes of these cell lines.

To determine whether the κ -transgenes had also undergone mutations in the cells where the H-chain gene was known to be mutated, genomic DNA libraries were produced from the hybridomas CE9, CF2 and 3A1. Fifteen clones each (about one for each copy, see Table 1) were analysed from CE9 and CF2; and four clones from 3A1 (which has 3 copies). Results of sequencing the V genes of these clones are shown in Fig. 2. In ~7.5 kb (~225 bp per clone) of V region sequence, four clones showed a total of 8 mutations. At least one mutated clone was derived from each cell line. Two clones, pIII-2 and pIII-4, bear one identical mutation; they probably represent the same gene (from a mouse with few transgenes) isolated twice. Two of these clones contain 3 mutations each, whereas the 30 other V genes sequenced in this area are identical to that of the original microinjected gene (this includes four MOPC-167 marker mutations, that differ from the endogenous germline V_κ gene, see asterisks in Fig. 2b). Part of the C' and 3' untranslated regions in the same clones were also sequenced; out of ~9.5 kb (~280 base pairs (bp) per clone) (Fig. 2a), no mutations were found. Thus, mutations in only the V region of the κ -transgenes were detected.

Some of the mutations introduce amino-acid changes (Fig. 2b). Note that the first mutation in pI-7 actually introduces a stop codon, and thus must inactivate the gene. Perhaps, because another mutation in the same gene causes the threonine in position 69 to be replaced by a much more hydrophobic alanine, the inactivation by the stop codon allowed the cell to be selected by antigen.

The pIII-2/pIII-4 mutation would change a thymidine to a guanosine in a cDNA copy of messenger RNA derived from the mutant 3A1 gene. To determine whether this change was evident in mRNA from this low transgene copy hybridoma, we sequenced V_κ-specific cDNAs prepared from mRNAs of the 3A1 hybridoma and of its identical twin, hybridoma 2F5. In both cases, co-expression of this mutation was detected along

with the unmutated base. As a negative control, we also checked four IgM-producing hybridomas derived from the same mouse; in these the guanosine was undetectable (data not shown). This is evidence that at least two of the three transgenes are transcribed. Comparing the relative intensity of the guanosine signal to that of the thymidine signal on the sequencing autoradiogram, we estimate that all three transgenes in 3A1 are transcribed.

The pIII-2/pIII-4 mutation described above (Fig. 2b) introduces a *TaqI* site into the DNA of one copy of the transgene. This shortens a 1,067-bp *TaqI* fragment to 1,002 bp, which is present only in this hybridoma, and not in kidney DNA of the mouse used to make the hybridoma or in another hybridoma derived from the same mouse (Fig. 3). Thus, the mutation must have occurred somatically in the original B cell giving rise to the hybridoma. Judging by the intensity of the 1,002-bp relative to the 1,067-bp band, only one of the three genes contained the mutation.

These experiments demonstrate that a rearrangement event is not required for immunoglobulin gene somatic hypermutation and are consistent with the accumulation model⁸. Indeed, it appears that the mutation process may only become activated in the course of the immune response, as others had postulated^{14,22}. We found it necessary to select hybridomas with mutations in the endogenous V_H gene in order to observe mutations in transgenes. In previous experiments not using antigen selection, we did not obtain evidence for somatic mutation of rearranged κ transgenes (ref. 23 and R.L.O'B. E. Selsing and U.S. unpublished data); we had isolated spleen cDNA clones of unselected transgene-encoded mRNAs from κ transgenic mice polyclonally immunized with sheep red blood cells. The cDNA clones that were clearly derived from the transgene, when sequenced, did not reveal mutations above the level that would be expected from errors introduced by reverse transcriptase (1–3 per 3,000 bp)^{24,25}. We found three base changes out of about 9,200 sequenced, in the C' region. In contrast, the data we present here reveal definite V-region-specific somatic mutations in a κ transgene; they differed from our preliminary studies by deliberate use of antigen selection.

The frequency of mutation in the transgene is relatively low. In the experiments reported here, the frequency (~1 base in every 900 sequenced) is four to five times lower than that seen

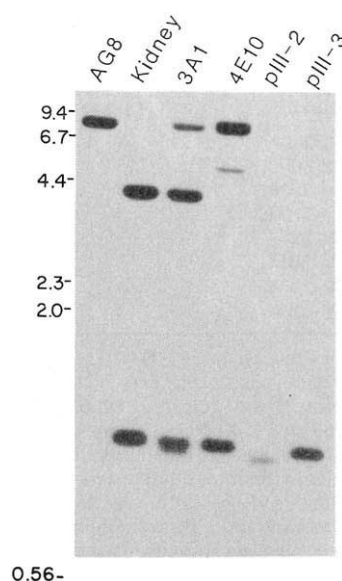


Fig. 3 Southern blot³⁸ of *TaqI*-digested DNAs. The hybridization probe represented part of *J_κ4*, *J_κ5* and part of the *J - C_κ* intron (*XbaI - Aval* fragment from a MOPC-21 κ clone). Size markers are to the left, in kb. The *TaqI* site created by mutation in one of the transgenes in 3A1 gives a 1,002-bp band in addition to the 1,067-bp band found in DNA of the mouse used to make the hybridomas (kidney). Only the 1,067-bp band is seen in another hybridoma from the same mouse (4E10). Lanes pIII-2 and pIII-3 represent 100–300 pg of cloned DNA for comparison; pIII-2 represents the mutated version of the 3A1 transgene. The top-most band is the rearranged κ gene of Ag8.653, the hybridoma fusing line. The next is the C57BL/6 germline κ band. Hybridoma 4E10 lacks the germline *J_κ* and contains a faint, high molecular mass band that is probably due to an endogenous *V_κ* rearrangement in a B cell lineage devoid of feedback inhibition³⁹.

(R.L.O.-B., E. Selsing and U.S. unpublished data). The lower frequency could reflect special conditions associated with a transgenic system. First, the chromosomal position of the transgene is different from the endogenous genes, which could have an effect, particularly if gene conversion-like recombination plays a role in the hypermutation process. Second, the transgene already contained somatic mutations derived from the original myeloma, and these might lessen the need or chance for the genes to mutate further to increase antigen affinity. Third, if we can generalize from the available evidence to conclude that all the multi-copy transgenes are transcriptionally active, B cells with mutated κ transgenes may not be selected by antigen. Fourth, having multiple genes in tandem might dilute out mutation enzymes and thus decrease the frequency. This appears a less likely explanation, because the frequency of mutations in the low and high copy hybridomas studied here is similar. Because the frequency is lower than that seen in hybridomas and myelomas from normal mice, the question arises whether the mutations might represent only ordinary, spontaneous somatic mutations not derived from immunoglobulin gene hypermutation. Mutation rates for V genes have been calculated, making various assumptions, as 10^{-4} per bp per cell generation¹⁴ to 10^{-3} per bp per cell generation¹⁵; these are several orders of magnitude greater than calculated spontaneous rates for other genes¹⁰. Although we did not attempt to calculate a mutation rate, two of the DNA clones shown here contain three mutations each; the probability of finding three spontaneous mutations in one clone would equal the spontaneous rate to the third power. Even allowing for antigenic selection it is difficult to believe we would have detected such an occurrence, especially twice. Moreover no mutations were detected in the C region or 3' untranslated region. Therefore, we conclude the mutations arise

in the same V_{κ} gene when present as a single endogenous copy in hybridomas and myelomas⁸, whereas mutations detected in the endogenous V_H genes in the same cells are comparable to levels typically seen for this V_H gene⁷. The lower frequency is not related to the choice of the V_{κ} gene as transgene, because in a preliminary experiment, a transgene representing the expressed κ -gene of MOPC-21, when isolated from a random IgG hybridoma, was also shown to be mutated at a low level

from an immunoglobulin gene hypermutation mechanism.

In addition to showing that ongoing gene rearrangement is not required for the hypermutation, the data presented here allow the following conclusions. (1) The chromosomal location of the gene is not crucial for somatic mutation. In fact, two different exogenous sites in the two transgenic lineages permit mutations. (2) The localization of the mutations to the *V* gene suggests that a special mutator exists that must recognize specific sequences in or around immunoglobulin *V* genes. Such sequences must then be present in the transgene used in these experiments. (3) The *C* region and 3' untranslated region were apparently not susceptible to hypermutation. This is in agreement with studies in normal mice^{6,7} and in contrast to the mutations sometimes seen in myeloma cell lines^{26,27}. Perhaps mutations in *C* genes are incompatible with normal antibody function and are therefore not present in B cells selected by antigen. But, in the experiments presented here, the κ transgenes exist in multiple copies so that *C* region mutations in one or a few transgenes may still permit the B cell to interact with antigen. Also, mutations were not found in the non-coding portion of the *C* exon. The absence of mutations near the *C* gene suggests the mutator is confined to the vicinity of the *V* region.

In conclusion, we have shown that an as yet unknown molecular mechanism is capable of locating and targeting mutations in a rearranged *V* gene by means of DNA sequences present within the transgene.

We thank Erik Selsing for many useful suggestions and for communicating related results, Sima Shabestari and Wendy Brown for help in phage and plasmid preparation, Roger Perlmutter for a *T₁₅V_H* primer and the unpublished *T₁₅V_H*-like sequence of the C57BL/10 mouse, Frank Burns for mRNA sequencing protocol, and Peter Engler and Jim Lewis for helpful advice and discussion. This work was supported by grants CA/AI 25754 and HD 23089 from the NIH to U.S., and to R.L.B. R.L.O. was supported by training grant PHS NRSA 5T32 GM7270 from NIGM.

Received 5 November; accepted 29 December 1986.

- Weigert, M. & Riblet, R. *Cold Spring Harb. Symp. quant. Biol.* **41**, 837-846 (1976).
- Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
- Nishioka, Y. & Leder, P. *J. biol. Chem.* **255**, 3691-3694 (1980).
- Selsing, E. & Storb, U. *Cell* **25**, 47-58 (1981).
- Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* **25**, 59-66 (1981).
- Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. *Cell* **27**, 573-581 (1981).
- Gearhart, P. J., Johnson, N., Douglas, R. & Hood, L. *Nature* **291**, 29-34 (1981).
- Gearhart, P. J. & Bogenhagen, D. T. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3439-3443 (1983).
- Bothwell, A. L. M. *et al. Cell* **24**, 625-637 (1981).
- Vogel, F. in *Chemical Mutagenesis in Mammals and Man* (eds Vogel, F. & Rohrborn, G.) 16-68 (Springer, Heidelberg, 1970).
- Gorski, J., Rollini, P. & Mach, B. *Science* **220**, 1179-1181 (1983).
- Sablitsky, F., Weisbanm, D. & Rajewsky, K. *EMBO J.* **4**, 3435-3437 (1985).
- Rudikoff, S., Pawlita, M., Pumphrey, J. & Heller, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2162-2166 (1984).
- Griffiths, G. M., Berek, C., Kaartinen, M. & Milstein, C. *Nature* **312**, 271-275 (1984).
- McKean, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3180-3184 (1984).
- Sablitsky, F., Wildmer, G. & Rajewsky, K. *EMBO J.* **4**, 345-350 (1985).
- Clarke, S. H., Clafflin, J. L. & Rudikoff, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3280-3284 (1982).
- Storb, U. *et al. J. exp. Med.* **164**, 627-741 (1986).
- Gershenson, H., Tsukamoto, A., Weissman, I. & Joho, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7674-7678 (1981).
- Perlmutter, R. M., Benson, B., Griffin, J. & Hood, L. *J. exp. Med.* **162**, 1998-2016 (1985).
- Brodeur, P. H. & Riblet, R. *Eur. J. Immunol.* **14**, 922-930 (1984).
- Wysocki, L., Manser, T. & Giffert, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1847-1851 (1986).
- O'Brien, R. L. thesis. Univ. Washington, Seattle (1986).
- Battula, N. & Loeb, L. A. *J. biol. Chem.* **250**, 4405-4409 (1975).
- Gerard, G. F. *BRL Focus* **8**, 12 (1986).
- Wahl, M., Burrows, P. D., von Gabain, A. & Steinberg, C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 479-482 (1985).
- Secher, D. S. *et al. Immun. Rev.* **36**, 51-72 (1977).
- Morrison, S. L. & Scharff, M. D. *CRC Crit. Rev. Immun.* **3**, 1-22 (1981).
- Kearney, J. F., Radbruch, A., Liesegang, B. & Rajewsky, K. *J. Immunol.* **123**, 1548 (1979).
- Nowinski, R. C. *Virology* **93**, 111-126 (1979).
- Clafflin, J. L., Clafflin, J., Hudak, S., & Maddalena, A. *J. exp. Med.* **153**, 352-364 (1981).
- White, B. A. & Bancroft, F. C. *J. biol. Chem.* **257**, 856-857 (1981).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-559 (1977).
- Kabat, E. A., Wu, T. T. & Bilofsky, H. *NIH Publ. No.* 80-2008 (Bethesda, Maryland, 1979).
- Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Viera, J. & Messing, J. *Gene* **19**, 259-268 (1982).
- Tobak, H. F. & Flavell, R. A. *Nucleic Acids Res.* **5**, 2321 (1978).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Gollahon, K. G., Brinster, R. & Storb, U. *J. exp. Med.* (submitted).

Rearrangement of a chicken immunoglobulin gene occurs in the lymphoid lineage of transgenic mice

D. Bucchini*, C-A. Reynaud†, M-A. Ripoché*, H. Grimal†, J. Jami* & J-C. Weill†

Institut Jacques Monod, * Unité INSERM 257 and

† Groupe d'Immunodifférenciation moléculaire, CNRS,

Université Paris VII, 2 Place Jussieu, 75251 Paris Cedex 05, France

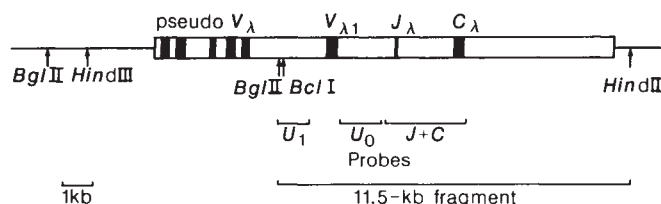
Immunoglobulin (Ig) and T-cell antigen receptor genes rearrange through identical heptamer-nonamer recognition sequences during entry of cells into the B or T lymphoid lineage¹⁻⁵. A similar enzymatic machinery may be used to perform these highly cell-specific events in these two types of lymphoid cells⁶. We have investigated what the signal may be that triggers the rearrangement of one or other of the receptor genes in B or T cells. Mice from three transgenic lines carrying two, four or twenty copies of the unrearranged chicken λ light-chain locus⁷ were analysed. In all three lines the chicken Ig transgene rearranges in B cells; in the line with 20 copies, a rearranged fragment can also be detected in thymus DNA. We conclude that the inserted chicken light-chain locus in its natural configuration contains target sequences that permit specific rearrangement in mouse lymphoid B cells, but that this precise differentiation step may be deregulated in thymic cells when the physiological level of relevant information is experimentally altered.

The microinjected DNA fragment (11.5 kilobases (kb)) encodes the chicken immunoglobulin λ light-chain locus in its germline configuration⁷. It contains the *V_λ1*, *J_λ* and *C_λ* coding exons, 1.4 kb 5' and 4.8 kb 3' of chicken genomic flanking regions (Fig. 1). Approximately 50 copies were microinjected into male pronuclei of F₂ eggs obtained after the mating of (C57BL/6 × CBA/J) F₁ mice. Here we present the results from the first generation offspring of three different transgenic mice.

The different mice analysed in each line show the same restriction pattern of foreign DNA indicating that there may be a single integration site in each case. A unique *Bcl*I restriction site is present in the 5' region of the transgene (Fig. 1). For the three lines, a *Bcl*I digest of liver DNA reveals, after hybridization with a (*J* + *C*)_λ probe (Fig. 1) an 11.5-kb fragment identical in size to the microinjected DNA, which suggests a head-to-tail tandem arrangement of the insertion. The three mouse lines have ~2, 4 and 20 copies of the inserted transgene (hereafter referred to as the 2-, 4- and 20-copies lines). Depending on the mouse line and on the probe used, another fragment, probably junctional, can be detected (Figs 2 and 3).

In chicken B-lymphoid tissues, rearrangement of the chicken λ light-chain locus involves the deletion of the intervening sequence between *V_λ* and *J_λ* and generates a restriction fragment 2 kb smaller than its germline counterpart⁷. The presence of such a restriction fragment of 9.5 kb is observed in the 4-copies transgenic mouse spleen DNA (Fig. 2h). This fragment represents 1-2% of the inserted chicken DNA. The rearranged fragment does not give any signal with a DNA probe for the deleted region (*U₀*) between *V* and *J* (Fig. 2f) but it hybridizes with both 5' (*U₁*) and 3' (*J* + *C*) probes (Fig. 2c, j). There is no such rearrangement in either non-lymphoid tissues (liver, brain) (Fig. 2a, d, k, l) or thymic cells (where the detection level is up to 20-fold greater than in spleen, so that rearrangement occurring in 0.1% of thymic cells would be detected) (Fig. 2b, e and g). When splenic B cells are enriched by anti- θ treatment, the rearranged fragment appears to be stronger (Fig. 2j), implying that splenic T cells do not markedly contribute to the rearrangement event observed. The λ light-chain phenotype is the major expressed isotype in the chicken ($\geq 90\%$)⁸, whereas it represents

Fig. 1 Map of the chicken immunoglobulin λ light-chain genomic region inserted in the C_λ 36 phage. Recombinant phages spanning the chicken light-chain locus were isolated from a genomic DNA library prepared from chicken erythrocytes (CB inbred line). The C_λ 36 phage inserted in the λ L47.1 vector (thin line) contains a 15-kb DNA fragment of the light-chain locus (open box). The 11.5-kb *Bgl*II–*Hind*III DNA fragment indicated contains the variable (V_λ 1), junctional (J_λ) and constant (C_λ) gene, 1.4 kb 5' and 4.8 kb 3' of flanking regions, and at the 3' end 500 base pairs (bp) of left-arm phage DNA. It does not include the several V_λ pseudogenes located upstream from V_λ 1. This fragment was isolated from the phage on a preparative 0.6% agarose gel and further purified by sucrose gradient centrifugation. After ethanol precipitation it was redissolved in 10 mM Tris pH 7.5, 0.25 mM EDTA and microinjected into pronuclei of fertilized mouse eggs. *Hind*III, *Bgl*II and *Bcl*I restriction sites are marked. The probes used in Figs 2 and 3 are indicated: U_1 , U_0 , ($J+C$).



only a minor population in the mouse (~5% of B cells). Transgenic mouse splenic B cells were sorted for the mouse λ^+ isotype and compared with the excluded λ^- fraction. The intensity of labelling of the rearranged chicken DNA fragment was almost identical in these two cell populations (Fig. 2c, i). This result does not imply a preferential pattern of rearrangement in cells having one or the other isotype.

In the 2-copies mouse line, rearrangement of only one of the copies of the transgene is observed (Fig. 3b), but both fragments obtained with a *Bcl*I digest contain a complete copy of the transgene, as demonstrated by various restriction enzymes. This rearrangement is quantitatively similar to the one observed in the 4-copies mouse line. No rearrangement can be detected in thymus or liver DNA (Fig. 3a, c) even when the detection level is increased 10- to 20-fold. In the 20-copies mouse line, a similar rearrangement is observed in spleen DNA (1–2% of the inserted transgene) (Fig. 3e). But, in contrast with the two other transgenic lines, rearrangement is also observed in thymic DNA and amounts to about one quarter of the rearrangement observed with spleen cells (Fig. 3f). No rearrangement of the transgene was observed in the liver (Fig. 3d). G1 mice of a fourth transgenic line containing ~20 copies of the transgene were also analysed. As in the 20-copies line, there was a rearrangement of the chicken transgene in spleen and thymus cells in similar proportions, and no rearrangement in non-lymphoid tissues (data not shown).

Rearrangement of the chicken immunoglobulin λ gene in the spleen of the 2-, 4- and 20-copies lines was further demonstrated by the presence of specific transcripts in the spleen (Fig. 4b, d and f). These transcripts are of normal size when compared with chicken spleen poly(A)⁺ RNA, and the amount of expression can be estimated, depending on the transgenic line, as 5–20% of the endogenous expression estimated with a mouse C_λ probe.

There was no expression of the chicken transgene in the thymus of the 2- and 4-copies lines (Fig. 4a, c). This held also for the 20-copies line in which rearrangement in thymus had been detected (Fig. 4e, where 0.5 μ g of thymus poly(A)⁺ RNA is analysed). On the Northern blots, a faint signal was observed in some cases with the chicken C_λ probe, but normalization with a mouse C_λ probe indicated that this expression might be due to contaminating parathymic lymph nodes⁹ (data not shown).

Rearrangement of the chicken λ light-chain gene observed in the thymus of the two 20-copies transgenic mouse lines was unexpected. A D–J joining at the Ig heavy-chain locus has been described in T cells¹⁰ but endogenous rearranged immunoglobulin light-chain genes in these cells have never been reported. The presence of twenty additional copies of V–J–C units coding for an immunoglobulin light-chain gene might deregulate this tissue-specific event in mouse thymic cells. Twenty additional copies of V genes will not increase significantly the endogenous light-chain V gene pool already present in mouse DNA and it is tempting to speculate that the J–C DNA region is a more likely candidate for containing the specific signals that trigger the rearrangement of the B-cell receptor in B cells.

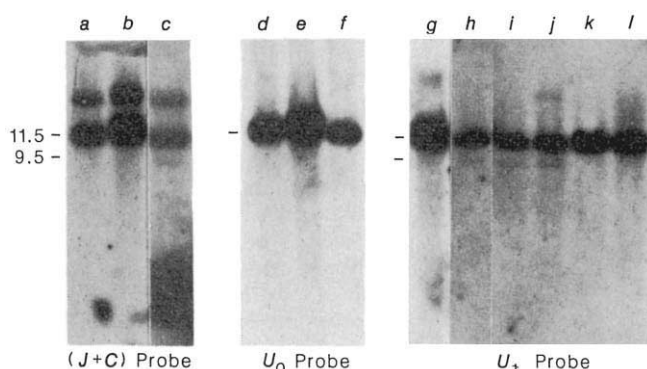


Fig. 2 DNA analysis of the 4-copies line. DNA from: a, d, k, liver (10 μ g); c, f, λ^- cells (5 and 10 μ g respectively); b, e, g, thymus (20 μ g); h, spleen (5 μ g); i, λ^+ cells (3–4 μ g); j, B cells (5 μ g); l, brain (10 μ g) (2 weeks film exposure). Nick-translated probes ($1-3 \times 10^6$ c.p.m. μ g⁻¹) were used as follows: lanes a–c, ($J+C$); lanes d–f, U_0 ; lanes g–l U_1 and were as described in Fig. 1. The band at 11.5 kb after digestion with *Bcl*I implies that the chicken DNA fragment was tandemly integrated in head-to-tail arrangement (~3–4 copies when compared with a single-copy gene). The band at 9.5 kb is the rearranged fragment. This band was already detectable after two days of exposure in spleen, B, λ^+ and λ^- cell populations, and does not hybridize with the U_0 probe. The band at 19 kb probably corresponds to a junctional fragment.

Methods. Three four-week-old transgenic females from the first generation progeny of the same transgenic mouse were examined. DNA was prepared from thymus spleen, liver and brain. Erythrocyte-free spleen cells were prepared (from one of the mice) as described¹⁵. B cells were obtained by treating spleen cell suspension with the monoclonal anti-Thy1.2¹⁶ and guinea-pig complement. The λ -positive cells were stained using the monoclonal rat anti-mouse $\lambda 1 + \lambda 2$ (L22 411)¹⁷ and the fluoresceinated mouse anti-rat light chain MAR 18¹⁸. Positive and negative cells were sorted using the Beckton Dickinson FACS4. The degree of purity of the λ^+ fraction was adjusted to obtain $\sim 10^6$ cells. This fraction was 72% λ^+ . DNA was prepared from these different cell preparations: B cells, λ^+ and λ^- fractions. Southern blot analysis of *Bcl*I-digested DNA from transgenic mice was performed as described^{19,20}.

The mice containing two copies of the transgene rearrange only one of them in splenic B cells. The simultaneous rearrangement of two adjacent sets of substrates could possibly be prevented in such a topological configuration, which does not occur in physiological conditions.

In spleen cells of the 2-, 4- and 20-copies lines, the respective values obtained for rearrangement and expression are of the same order of magnitude, suggesting that once the chicken transgene is rearranged, it is expressed normally. In contrast, transcripts of the chicken gene were not detected in thymus, even in the 20-copies line where the transgene was rearranged. Absence of expression of a rearranged immunoglobulin light-chain transgene in thymic cells even when in high copy number has been reported¹¹.

Heptamer-nonamer signals have been shown to be operative throughout the evolutionary scale^{12,13} and the recombinase machinery may have been conserved as well¹⁴. It remains to be

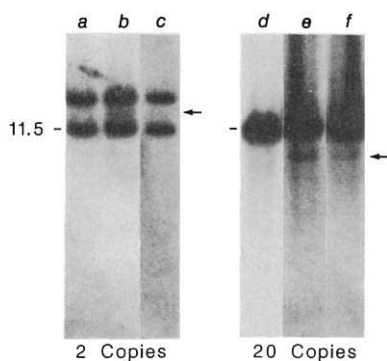


Fig. 3 DNA analysis of the 2-copy and 20-copy lines. *BclI*-digested DNA (10 µg) from liver (a, d), spleen (b, e), thymus (c, f) of the 2-copy (a-c, left panel) and 20-copy (d-f, right panels) mouse lines were analysed by Southern blotting and hybridization with a $(J+C)_A$ probe (conditions as in Fig. 2). Migration position of the 11.5-kb *BclI* fragment is indicated (compare Fig. 2). Arrows, rearranged fragments (3 days exposure). No rearrangement can be detected in the thymus of the 2-copy line after 3 weeks exposure. These fragments do not hybridize with the U_0 probe (not shown).

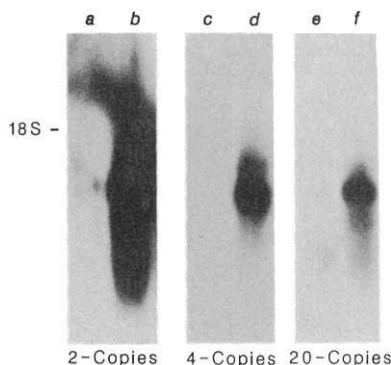


Fig. 4 Expression of the chicken transgene in lymphoid tissues. a, b, Thymus total RNA (4 µg) and spleen poly(A)⁺ RNA (5.5 µg) from the 2-copy line; c, d, thymus total RNA (10 µg) and spleen poly(A)⁺ RNA (1 µg) from the 4-copy line; e, f, thymus and spleen poly(A)⁺ RNA (0.5 µg and 0.2 µg respectively) from the 20-copy line; all probed with a chicken C_A probe (see below). Migration position of 18S ribosomal marker is indicated. The size of the chicken transcript corresponds to the normal size of the chicken λ light-chain mRNA (not shown).

Methods. RNA was extracted by the citric acid technique and poly(A)⁺ RNA isolated by oligo-dT cellulose chromatography²¹. RNA samples were fractionated through formaldehyde-agarose gels, transferred to nitrocellulose filters²² and hybridized with a chicken C_A probe: a 590-bp *HpaII* fragment from a chicken cDNA plasmid²¹ containing the entire constant and 3' untranslated region. The probe was labelled by oligo-labelling²³ to 4×10^9 c.p.m. µg⁻¹. Hybridization conditions were as in Fig. 2 (20-h film exposure). To normalize for the endogenous mouse κ expression, the filters were dehybridized and probed in the same conditions with a mouse C_K probe: a 1.95-kb *HhaI* fragment containing the entire insert of the $P_K(II)24$ plasmid ($V_{H9}+C+3'$ untranslated region: 900 bp) which was a gift of Dr R. P. Perry²⁴ (data not shown).

explained why the rearrangement of the chicken Ig transgene is found at such a low level. This may be caused by the sites of insertion of the transgene, or else, by negative selection acting on the cells where the rearrangement has occurred. Alternatively, the specific signals regulating the recombination of Ig genes may have diverged between birds and mammals, whereas the signals that control the expression of these genes have apparently remained closely related.

We thank Dr M. H. Julius and the staff of the Basel Institute for Immunology for making the cell sorting possible, and Mrs

C. Sagot for typing the manuscript. The technical assistance of V. Anquez and P. Lorès is gratefully acknowledged. This work was supported by funds from the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique, the Fondation de France and the Association pour la Recherche sur le Cancer.

Received 14 November; accepted 22 December 1986.

1. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
2. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
3. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981-992 (1980).
4. Chien, Y. H., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
5. Siu, G. *et al. Cell* **37**, 393-401 (1984).
6. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L. & Alt, F. W. *Cell* **44**, 251-259 (1986).
7. Reynaud, C.-A., Anquez, V., Dahan, A. & Weill, J.-C. *Cell* **40**, 283-291 (1985).
8. Grant, J. A., Sanders, B. & Hood, L. *Biochemistry* **10**, 3123-3132 (1971).
9. Grosschedl, R., Weaver, D., Baltimore, D. & Costantini, F. *Cell* **38**, 647-658 (1984).
10. Kurosawa, Y. *et al. Nature* **290**, 565-570 (1981).
11. Storb, U., O'Brien, R. L., McMullen, M. D., Gallahan, K. A. & Brinster, R. L. *Nature* **310**, 238-241 (1984).
12. Litman, G. W. *et al. Nature* **303**, 349-352 (1983).
13. Litman, G. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2082-2086 (1985).
14. Hope, T. J., Aguilera, R. J., Minie, M. E. & Sakano, H. *Science* **231**, 1141-1145 (1986).
15. Julius, M. H. & Herzenberg, L. A. *J. exp. Med.* **140**, 904-920 (1974).
16. Marshak-Rothstein, A. *et al. J. Immun.* **122**, 2491-2497 (1978).
17. Weiss, S., Lehman, K. & Cohn, M. *Hybridoma* **2**, 49-54 (1983).
18. Lanir, L. L., Crilman, C. A., Lewis, D. E., Griswold, S. T. & Warner, N. L. *Hybridoma* **1**, 125-131 (1982).
19. Southern, E. *J. molec. Biol.* **98**, 503-517 (1975).
20. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
21. Reynaud, C. A., Dahan, A. & Weill, J. C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4099-4103 (1983).
22. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271-281 (1985).
23. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
24. Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15**, 1495-1509 (1978).

Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes

C. R. Vossbrinck*,†, J. V. Maddox*, S. Friedman†, B. A. Debrunner-Vossbrinck‡ & C. R. Woese‡

* Office of Agricultural Entomology and Illinois Natural History Survey, 607 E. Peabody, Champaign, Illinois 61820, USA

† Department of Entomology, 505 S. Goodwin, University of Illinois, Urbana, Illinois 61801, USA

‡ Department of Microbiology, 407 S. Goodwin, University of Illinois, Urbana, Illinois 61801, USA

The microsporidia are a group of unusual, obligately parasitic protists that infect a great variety of other eukaryotes, including vertebrates, arthropods, molluscs, annelids, nematodes, cnidaria and even various ciliates, myxosporidia and gregarines¹. They possess a number of unusual cytological and molecular characteristics. Their nuclear division is considered to be primitive², they have no mitochondria³, their ribosomes and ribosomal RNAs are reported to be of prokaryotic size^{4,5} and their large ribosomal subunit contains no 5.8S rRNA⁶. The uniqueness of the microsporidia may reflect their phylogenetic position, because comparative sequence analysis shows that the small subunit rRNA of the microsporidium *Vairimorpha necatrix*^{7,8} is more unlike those of other eukaryotes than any known eukaryote 18S rRNA sequence. We conclude that the lineage leading to microsporidia branched very early from that leading to other eukaryotes.

Although higher taxa of eukaryotes, such as kingdoms and phyla, have been generally defined by macro- and micro-morphological criteria, opinions in the field of protist taxonomy vary widely, making the understanding of evolutionary relationships in this group uncertain (for reviews see refs 9-12). Recently developed molecular sequencing methods, which have already been used in establishing prokaryotic phylogenies^{13,14}, may be useful in resolving some of these uncertainties. From initial work¹⁵ we believe a relatively reliable and interesting picture of

(1-100)		
<i>S. cerevisiae</i>	-----U.U.UC..	...A.UA..C AU.U...UG. .U.A..... .G..... .-.C.AA.U AU----- --AAGCAAU UAUACAG..A
<i>V. necatrix</i>	-----CACCA GGUUGAUUCU	GCCUGACGUA -GACGCUAUU CCCUAGAUAU ACCCAUGCAU G-UUUUUA UA----- --UGG
<i>E. coli</i>	AAAUUG..AG. .U.....CA.	.G..C.GA.U GA....GGC GG.AGG-CC.. .A.....A .UCGAACG.U A.CAGGAAGA AGCUUGCUUC UUUGCUGAC.
(101-200)		
<i>S. cerevisiae</i>	..CUCC.A.UG....U... AU.AUGUA.CG...U.UG. AGUUCUUUA**A.G.... .C.U.G.A.UUC.AG...U .AU.C.UGCUUAAAAUCUGC *****	
<i>V. necatrix</i>	AAAAUGGACUGCUCAGUAU	UACUCACUUUAUUUAUGUAU UAAA-----UUAGUAU AACUGCGUU-AAAGUGUAGC- AUAAGACAU-----
<i>E. coli</i>	.GUGCG....G.G.G.....	.GUCUGGGA-.AC.GCCU.. GG.-----GGG.G.. .A.UGG-.CG....U .AU.CCGCAUAAACGUCGCAA GACCAAAGAG
(201-300)		
<i>S. cerevisiae</i>	CAUGGCCUUG UGCUGCGCAU	GGU.CAUUC. .AUU.CU.C.A..U UC.A..G..G.A .G....C.. C..U..UUUC A..... .GGAUAAG-
<i>V. necatrix</i>	-----	---AUACAGUA AGAGUGAGAC CUAUCAGCUA --GUUGUUAAGG UAAUGGCUUA ACAAGGCAU GACGGGUAACG GUUUAUCUUG
<i>E. coli</i>	GGGACCUUC GGGCCUCUUG	CC..CGGA.G U.CCCAGAU GG..U.... --.A.G.GG.. .C....C. C.U...G.C .UCCC..-GC UGG.CUGAG-A
(301-400)		
<i>S. cerevisiae</i>	GUUGCA.U..G.....	A..C.CA .C.....AGC. .C..A..AC .A..CCUA- A..CAGC... U...G.CAA
<i>V. necatrix</i>	UAAUAUU-CC GGACAGAG	CCUGAGAGAC GGCUAUAG UCUAAGCAUU GCACGAGGC CGAAACUUA CCUAUGCAU- UUUUAUCUGAG GCAGUAUUG
<i>E. coli</i>	GG..GACCAG CC.C.U.GA	A.....C.. .UCCAG.CU C...C..GAGU.. G...UA..C A.A....GCG CAAGC....UCC...C
(401-500)		
<i>S. cerevisiae</i>	U..A.AACGA U.C.GGGCCC-A UUC--GGGUC UUGUAU.UGG AA.GAG..CA AU-----	-----GUAU .UACC.UAAC .AG..AC...
<i>V. necatrix</i>	GAAGU----- AAUAUU-----	-----CUAU GUUUAUAUU -----GUAU AAGUAUAUGA GGUGUAUUAU
<i>E. coli</i>	CGC..GUAUG .G.A--GGCCU	UCGGUUGUA AAGUA..U.C ACCGGCG.CG AAGCGCAUAA AGUUAUACC UUUGCUCAUU G.CGU..CCC .CA..AG..G
(501-600)		
<i>S. cerevisiae</i>	G ..UG.....	U.CA.CU..C. .A...U.AAG ...U..... .CU.G.AC *****
<i>V. necatrix</i>	UGGAGGGCAA UCAAGUGCCA	GCAGCCGCGG UAAUACUUGU UCCAAGAGUG UGUUAUGAUA UUGAUGCAGU UAAAAAGUCC GUAGUUUAUA
<i>E. coli</i>	CACC-...U.. CUCC.....	GGAG GGUGCA..C. .UA..CGA. .AC..GGCGGC.CA. .C..CGCG.U UGUUAAGUA
(601-700)		
<i>S. cerevisiae</i>	*****	***** A.UAGG.ACG GUC.G..GCA U.G...UCA AUU.UC-... G.....UCU
<i>V. necatrix</i>	*****	-----UU UAAAGACAA UAUGAGGUGU ACUGUAUAGU UGGGACAGAG AUGAAUUGU
<i>E. coli</i>	GAUGUGAAU CCCCGGCGUC	AACCUGGGA CUGCAUCUGA UACUGGCAAG CUU..GU.UC GUA...G.G GUA.A.UCC A..UGU..C. G.....C.
(701-800)		
<i>S. cerevisiae</i>	UG..UUUAUU GAA...U...	UACU..... .GU.UGC.A A.G.C..U.. CG..A.U... .A...A...U .AG...U.AC...C.G.C.
<i>V. necatrix</i>	ACGACCCUGA CUGGACGAAC	AGAAGCGAAA GCUGUACACU UCUAUGUAUU UUUUGAACAA GGACGUAAGC UGGAGGACCG AAGAUGAUAU GAUACCAUUG
<i>E. coli</i>	UA..GAUCUG GA...AU.C.	G..UG.....G .G.CC.C.. G.ACCAAGAC .GAC.CU..G .UG..A... GU.G....A .C.G..... .C.G.
(801-900)		
<i>S. cerevisiae</i>	CUU.A. CA.....	U.GA .CGGG..G.G **CCAC.C. G-ACC.UAC G.....CAA ...C...G..U.GG..G.....
<i>V. necatrix</i>	UAGUCCAGC AGUAAACUAU	GCCGACGAUG UGAUAUGAUA -UUA--AUUG UAUUAUGAUA UAGAAUUUG AGUUUUUUGGU CUGGGGAUAG UAUAGCGCA
<i>E. coli</i>	C.AC.. C.....G..	.U...UUG. A.G.UGGCC C..GAGGCGU GGC.UCCG.. GCU..CGCGU UAAG--C.A.C GCCU..GG.. .C.GC...
(901-1000)		
<i>S. cerevisiae</i>	..GC....CG...	G GC...CU...GCC. A..G..A... .C...GGUC C----AG.C
<i>V. necatrix</i>	AGAUGGAAA UUAAGAAAU	UGACGGAAGA AUACCAAGAAG AGUGGAUUGU CGCGCUUAU UUGACUAAAC GCGAGGUAAU UUAACAAUUA U-----UUAU
<i>E. coli</i>	..G..A..C C...UG..	GG.C CCG.-C..C G.....GCA. .U..U.... .C..UG.... .A.A.C.UGGUC .UGACA.CCA
(1001-1100)		
<i>S. cerevisiae</i>	A..-AUA.G.GA.AGAU.GCUCUU UCU...UUU	GUGGG.G...C...G..U .G..GA.... UU.G.C..C.G.GAU
<i>V. necatrix</i>	UCA-GAAGAG AUUUUGAUCU	GAGAA-----UGAU-- --AAUAGUG GUGCAUGGCC GUUUUCAUG GAUGGUGUGA AGU-UUUGAU UAAUUUACCC
<i>E. coli</i>	CGGAAGUUUU CAGA-GAUGAG	A.UGUGCCUU CGGGA.CGU GAG.CAG... C.....U .CG..GCU CG..U.... .A.G..G.G. .G.C.CG.
(1101-1200)		
<i>S. cerevisiae</i>	..CGAAC... ..U.AACC. ACUAAAUAGU	GGUG**GUUA UCCACUUCUU AG.CG..UA U.GGU-UUCA. GCC.AU... .UUUGA.GCA .U.....
<i>V. necatrix</i>	AAGACCGUGAG ACCUUUUUAU	U-----AAUAGACAG ACAC--AAUA -GUGUAGGAA GGAAAGGAUU AAAACAGGUC
<i>E. coli</i>	..CGA.C.CAA.CC	.UUUGUGCCA CGGUCCGGC CGGGAACUCA A.GG...U. C..GU-G..A. AC..G.... .UGG...G .CGU..A..
(1201-1300)		
<i>S. cerevisiae</i>	U..G..... U...G..C..	C..... .GC....C. GACGGAGCCA GCGAG**CUU GGCCG.G.GG UCU.GGUAU
<i>V. necatrix</i>	CGUUAUGCCCU CAGACAUUUG	GGCUGGACGC GCAUAUCAAU -AGAUAUAUA AUCUU----- CUUGUGAAAC UCCGUGGUG
<i>E. coli</i>	-A.C...G..C UU-.GACCA.	A...A .UCC..... GCGCAUACA AAGAGAACCG ACCUCCGCGAG .G.AAGCGGA CCUCAUAAAC UGCGUCCUAG
(1301-1400)		
<i>S. cerevisiae</i>	UG.....GA GCA.....U	UAUUGCUC.. C...GA... .C..... GCGCA.G.C UC.GC.U.CGU.AC..CC..
<i>V. necatrix</i>	AUGGGAUAA AUUUUGUAU-	-GAGAUUAUU GAACUUGGAA UUGCUAGUAA AUUUUAUUA AUAAGUAGAA UUGAUUGUGU CCCUGUUCUU UGUACACACC
<i>E. coli</i>	UCC....UGG .G.C..C..C	UCGAC.CCA. .G.C.C... .C..... UCG.GGA.C. GA.U.CCAGC G.....AC. U..C.GC.C.
(1401-1500)		
<i>S. cerevisiae</i>	GUAACC..UG	.AUG.CU.A. ...GGCC.CA GGAUCUGC.* *CGGAGAAU U.GGACA... UUGG.CAU.. G..GGAAC.. A.....
<i>V. necatrix</i>	GCCCGGCGCU AUCUAAGAUG-	AUAUGUGUUG UGA-AAUUA- -UACUUGAAC AAUAUGUAU AGAUCUGAUA UAAGUCGUA
<i>E. coli</i>	A.A CCA.GG..GU-	GGG.UGCAA A..AGUAGGU AGCUUAACCU UCGGGAGGC GCUUACC.CU UUGUGA.UCA U..CUG.GGU G.....
(1501-1542)		
<i>S. cerevisiae</i>	..A...U.C ..A..U...	.UGCG.A... ..U.-- --
<i>V. necatrix</i>	CAUGGUUGCU GUUGGAGAAC	CAUUAGCAG AUCAUAA-- --
<i>E. coli</i>	..A...AA.C ..A.G....	.UGCG.UU..CCUCCU UA

Fig. 1 Alignment of the *Vairimorpha necatrix* small subunit rRNA sequence with those of *Saccharomyces cerevisiae*²⁰ and *Escherichia coli*²¹. The alignment procedure was as in ref. 22: those positions in the other two sequences that have the same composition as the *V. necatrix* sequence are shown as dots; dashes signify that no base occurs at the given position; *, regions of the *S. cerevisiae* sequence that have been omitted for convenience of presentation. Fragments produced by cleavage of *V. necatrix* DNA with restriction endonuclease *Bgl*II were inserted into the *Bam*HI site of phage λ L47³⁶. A clone carrying the small subunit rRNA gene was selected by plaque hybridization procedures, using ³²P-labelled *V. necatrix* small subunit rRNA as a probe. A 2.2-kilobase (kb) *Cl*aI fragment containing the entire gene was subcloned into the *Acc*I site of phage M13mp8 in both orientations³⁷, and the single-stranded phage DNAs used as templates for dideoxy sequencing of the rRNA gene³⁷⁻³⁹. Both the usual M13 primer and primers specific for ribosomal RNA genes were used³⁷. Ninety-five per cent of the sequenced portion of the rRNA gene was determined for both DNA strands. Numbering, spacing according to the standard *E. coli* sequence²¹.

eukaryotic phylogeny is beginning to emerge. Based on the limited number of complete, eukaryotic, small subunit rRNAs (18S) which have been sequenced, the animals, higher plants, certain fungi and some of the protists all seem to have arisen within a relatively short interval of time. It has also been shown

by using this method that two species of mastigote protists, *Trypanosoma brucei* and *Euglena gracilis*, at present treated by some protozoologists as members of a single phylum (Euglenozoa)¹⁰, in fact arose much earlier than this large plant-animal-fungal radiation.

Table 1 Pairwise percentage differences for various eukaryotic 16S-like rRNA sequences

Species	<i>V. necatrix</i>	<i>E. gracilis</i>	<i>T. brucei</i>	<i>D. discoideum</i>	<i>P. tetraurelia</i>	<i>S. cerevisiae</i>	<i>Z. mays</i>	<i>X. laevis</i>	<i>E. coli</i>
<i>V. necatrix</i>	—	63.6	65.4	56.3	57.0	56.6	58.8	61.1	—
<i>E. gracilis</i>	63.1	—	36.7	41.4	35.5	35.3	37.0	40.3	—
<i>T. brucei</i>	64.1	35.8	—	38.8	36.0	37.3	36.7	38.8	—
<i>D. discoideum</i>	54.8	40.6	37.3	—	22.5	22.7	22.7	27.8	—
<i>P. tetraurelia</i>	55.4	34.4	35.2	21.9	—	14.1	14.5	21.4	—
<i>S. cerevisiae</i>	55.4	34.8	36.2	22.2	13.4	—	12.5	18.2	—
<i>Z. mays</i>	57.6	36.4	35.8	22.2	14.2	12.1	—	14.8	—
<i>X. laevis</i>	59.7	39.5	37.9	27.1	20.8	18.1	14.4	—	—
<i>E. coli</i>	82.3	77.2	77.5	78.5	77.7	80.0	80.2	80.7	—
<i>S. solfataricus</i>	76.7	61.5	70.2	65.3	65.5	65.7	68.5	67.8	48.9

The set includes small subunit rRNA sequences of *Saccharomyces cerevisiae*²⁰ and *Escherichia coli*²¹ as well as *Euglena gracilis*¹⁵, *Trypanosoma brucei*¹⁵, *Dictyostelium discoideum*³⁰, *Paramecium tetraurelia*³¹, *Zea mays*³², *Xenopus laevis*³³ and *Sulfolobus solfataricus*³⁴. Per cent differences have been converted to evolutionary distances by the correction of Jukes and Cantor³⁵. Only those positions represented in all the sequences considered are scored. Upper right section (above diagonal): sequence set includes only the eukaryotic sequences indicated; 1,175 positions scored. Lower left section: sequence set includes, in addition, the two prokaryotic sequences; 1,114 positions scored.

There have been speculations concerning the phylogeny of the microsporidia for about 100 years. They became a high-ranking taxon when Balbiani¹⁶ proposed that microsporidia should be considered an order within the class Sporozoa. In the widely used classification of the Protozoa proposed by Wenyon¹⁷ the microsporidia were considered to be one of three orders in the class Cnidosporidia. (The other two orders in Wenyon's Cnidosporidia, the Myxosporidia and the Actinomyxidia, although having superficial resemblance to the microsporidia, are now not believed to be closely related to them¹⁸). The microsporidia were later elevated to the class Microsporea¹⁹ and finally to the phylum Microspora¹. The phylum was defined as "Protists that form unicellular spores, each with a polar filament or rudiment thereof. Obligate intracellular parasites. Without mitochondria."¹ Because of the unusual molecular and cytological characteristics of microsporidia we decided to sequence their small subunit rRNA to clarify their phylogenetic position and

otic small subunit rRNA, and (at only 1,244 nucleotides) even appreciably shorter than its prokaryotic counterpart. Moreover, the *V. necatrix* sequence lacks various regions of the molecule considered to be 'eukaryotic'. Note particularly the regions between positions 180 and 225, and positions 590–650 where little or no homology exists between eukaryotic and prokaryotic sequences^{22,23}; the *V. necatrix* sequence lacks both regions.

These characteristics of the *V. necatrix* small subunit rRNA gene are so unusual that a portion of the small subunit rRNA was sequenced directly^{24,25} to make sure that the gene sequence actually represents the functional form of the molecule. The regions sequenced extend from the 5' end of the molecule to position 302 and between positions 968 and 1,490 (*E. coli* numbering). No discrepancies between the rRNA and corresponding gene sequences were encountered.

There is little overall homology between the *V. necatrix* rRNA sequence and those of other eukaryotes (Table 1). Figure 2, a phylogenetic tree derived from these data by a distance-matrix tree-generating algorithm^{26,27}, shows that *V. necatrix* sequence is consistent with an early branch in the eukaryotic line of descent.

The fact that the root of the eukaryotic tree lies between the *V. necatrix* branch and all other eukaryotic lineages for which sequences are available also emerges from parsimony analysis²⁸. In fact, considering all possible combinations of four sequences, one of which is an outgroup (either eubacterial or archaeobacterial), the second of which is the *V. necatrix* sequence, and the remaining two are any other combination of eukaryotic sequences used in Table 1, the most parsimonious branching arrangement in all cases places the *V. necatrix* sequence with the outgroup sequence (analysis not shown).

Using the criterion of rRNA difference as a measure of evolutionary distance, the depth of the microsporidian branching in the eukaryotic line is remarkable. This lends credence to the idea that some of the organism's unique characteristics may indeed signify a split from other eukaryotes very early in time. Perhaps this divergence occurred when the Earth's atmosphere lacked free oxygen (a time according to geological estimates, around $2.9\text{--}2.7 \times 10^9$ yr BP; ref. 29), before the evolution of mitochondria, and before development of the 5.8S rRNA in the eukaryotes. Later, as O₂ built up in the atmosphere it is possible that of the life forms that had not evolved mitochondria, one group which survived was that which had become intracellular parasites of organisms containing mitochondria. As parasites they would not be subject to the effects of the oxygen-rich atmosphere and might avoid competition from organisms which contained mitochondria. Indeed, the fact that microsporidia parasitize such a broad range of animals and protists allows the possibility that the origin of their parasitism is very ancient.

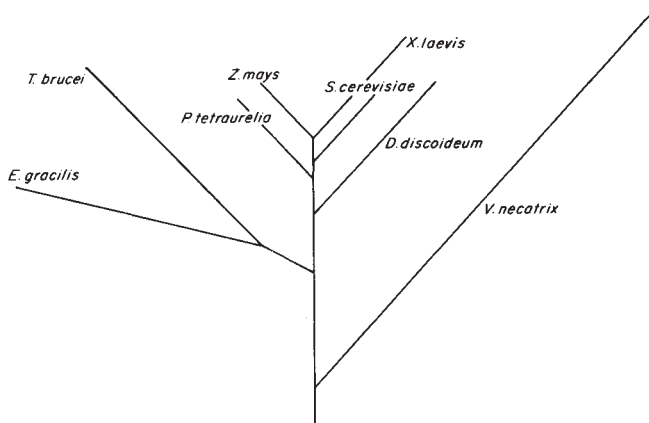


Fig. 2 Phylogenetic tree for the eight eukaryotes of Table 1 derived by a distance-matrix-treeing method²⁶. The root of the tree has been imposed subsequently by invoking the prokaryotic outgroups. Line length is proportional to evolutionary distance in Table 1.

thus test the hypothesis that these characteristics may be primitive rather than parasitically derived.

Figure 1 shows the microsporidia sequence aligned with a eukaryotic (*Saccharomyces cerevisiae*)²⁰ and a eubacterial (*Escherichia coli*)²¹ sequence. Table 1 lists pairwise percentage differences and evolutionary distances among these and other representative eukaryotic rRNA sequences.

The *V. necatrix* sequence is far shorter than a typical eukary-

This work was supported by a grant from NASA to C.R.W. We thank Drs G. DeSoete and D. Swofford for giving us their tree-generating algorithms. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00266.

Received 30 December 1986; accepted 3 February 1987.

1. Sprague, V. in *Comparative Pathobiology* Vol. 2 (eds Bulla, L. A. & Cheng, T. C.) 335-385 (Plenum, New York, 1977).
2. Raikov, I. B. *The Protozoan Nucleus: Morphology and Evolution*, 124-129 (Springer, Berlin, 1982).
3. Vavra, J. in *Comparative Pathobiology* Vol. 1 (eds Bulla, L. A. & Cheng, T. C.) 1-86 (Plenum, New York, 1976).
4. Ishihara, R. & Hayashi, Y. *J. Invert. Pathobiol.* **11**, 377-385 (1968).
5. Curgy, J. J., Vavra, J. & Vivares, C. *Biol. Cellulaire* **38**, 49-52 (1980).
6. Vossbrinck, C. R. & Woese, C. R. *Nature* **320**, 287-288 (1986).
7. Krammer, J. P. *Invert. Pathol.* **7**, 117-121 (1965).
8. Maddox, J. V. & Sprekel, R. K. *Misc. Publ. ent. Soc. Am.* **11**, 65-84 (1978).
9. Taylor, F. J. R. *BioSystems* **10**, 67-89 (1978).
10. Cavalier-Smith, T. *BioSystems* **14**, 461-481 (1981).
11. Lee, J. J., Hutner, S. H. & Bovee, E. C. *An Illustrated Guide to the Protozoa* (Society of Protozoologists, Lawrence, Kansas, 1985).
12. Margulis, L. *Symbiosis in Cell Evolution* (Freeman, San Francisco, 1981).
13. Woese, C. R. & Fox, G. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5088-5090 (1977).
14. Fox, G. E. *et al. Science* **209**, 457-464 (1980).

15. Sogin, M. L., Elwood, H. J. & Gunderson, J. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1383-1387 (1986).
16. Balbiani, E. G. *C. r. hebd. Acad. Sci., Paris* **95**, 1168-1171 (1882).
17. Wenyon, C. M. *Protozoology* (Bailliere, Tindall and Cox, London, 1926).
18. Vavra, J. *Proc. First Int. Congr. Parasitol.*, 443-444 (1964).
19. Corliss, J. O. & Levine, M. D. *J. Protozool.* **10** (suppl.), 26-27 (1963).
20. Rubtsov, P. M. *et al. Nucleic Acids Res.* **8**, 5779-5793 (1980).
21. Brosius, J., Palmer, J. L., Kennedy, J. P. & Noller, H. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4801-4805 (1978).
22. Woese, C. R., Gutell, R. R., Gupta, R. & Noller, H. F. *Microbiol. Rev.* **47**, 621-669 (1983).
23. Gutell, R. R., Weiser, B., Woese, C. R. & Noller, H. F. *Prog. nucleic Acid Res.* **32**, 155-216 (1985).
24. Qu, L. H., Michot, B. & Bachellerie, J. P. *Nucleic Acids Res.* **11**, 5903-5920 (1983).
25. Lane, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 6955-6959 (1985).
26. De Soete, G. *Psychometrika* **48**, 621-626 (1983).
27. De Soete, G. *Z. Naturforsch.* **38**, 156-158 (1983).
28. Swofford, D. *PAUP Users Manual Version 2.3* (Natural History Survey, Champaign Illinois, 1985).
29. Walker, C. G. *et al. in Earth's Earliest Biosphere* (ed. Schopf, J. W.) 260-289 (Princeton University Press, Princeton, 1983).
30. McCarroll, R. *et al. Biochemistry* **22**, 5858-5868 (1983).
31. Sogin, M. & Elwood, H. J. *J. molec. Evol.* **23**, 53-60 (1986).
32. Messing, J., Carlson, J., Hagen, G., Rubenstein, I. & Oleson, A. *DNA* **3**, 31-40 (1984).
33. Salim, M. & Maden, E. H. *Nature* **291**, 205-208 (1981).
34. Olsen, G. J. *et al. J. molec. Evol.* **22**, 301-307 (1985).
35. Jukes, T. H. & Cantor, C. R. in *Mammalian Protein Metabolism* (ed. Munro, H. N.) 21-132 (Academic, New York, 1969).
36. Loenen, W. A. M. & Brammar, W. J. *Gene* **20**, 249-259 (1980).
37. Messing, J. *Meth. Enzym.* **101**, 20-78 (1983).
38. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
39. Mills, D. R. & Kramer, F. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2232-2235 (1979).

Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA

Steven J. Brill, Stephen DiNardo*,
Karen Voelkel-Meiman & Rolf Sternglanz†

Department of Biochemistry, State University of New York,
Stony Brook, New York 11794, USA

Yeast strains with mutations in the genes for DNA topoisomerases I and II have been identified previously in both *Saccharomyces cerevisiae*¹⁻³ and *Schizosaccharomyces pombe*⁴. The topoisomerase II mutants (*top2*) are conditional-lethal temperature-sensitive (*ts*) mutants. They are defective in the termination of DNA replication and the segregation of daughter chromosomes²⁻⁵, but otherwise appear to replicate and transcribe DNA normally. Topoisomerase I mutants (*top1*), including strains with null mutations are viable and exhibit no obvious growth defects, demonstrating that DNA topoisomerase I is not essential for viability in yeast^{1,4,6,7}. In contrast to the single mutants, *top1 top2 ts* double mutants from both *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* grow poorly at the permissive temperature and stop growth rapidly at the non-permissive temperature^{4,7}. Here we report that DNA and ribosomal RNA synthesis are drastically inhibited in an *S. cerevisiae top1 top2 ts* double mutant at the restrictive temperature, but that the rate of poly(A)⁺ RNA synthesis is reduced only about threefold and transfer RNA synthesis remains relatively normal. The results suggest that DNA replication and at least ribosomal RNA synthesis require an active topoisomerase, presumably to act as a swivel to relieve torsional stress, and that either topoisomerase can perform the required function (except in termination of DNA replication where topoisomerase II is required).

It had been shown that a *top2 ts* mutant is capable of undergoing one, and only one, complete round of DNA replication at the non-permissive temperature². As a *top1* mutant grows normally, it obviously has no significant defect in DNA replication^{6,7}. In contrast, DNA replication is severely affected in a *top1 top2 ts* double mutant. Figure 1 shows the results of an experiment with a culture synchronized in the G1 phase of the

cell cycle by treatment with the yeast pheromone, α -factor. When cells are released from the α -factor block at the permissive temperature, 25 °C, they progress through the cell cycle, albeit slowly, and DNA synthesis during S phase is normal. On the other hand, if the cells are released from the G1 block at the non-permissive temperature, 37 °C, no significant DNA synthesis is seen. This does not by itself indicate a specific block in DNA replication, as the double mutant arrests in G1 with this protocol (ref. 7 and our unpublished data). A similar result is found when cells are shifted to 37 °C just at the beginning of S phase. Moreover, cultures shifted to 37 °C at any time during S phase cease DNA synthesis rapidly (Fig. 1). The results suggest that topoisomerase activity is required continuously for continuing DNA replication and that either topoisomerase can perform the necessary function. Only when the activity of both enzymes is abolished, in the double mutant at 37 °C, does DNA synthesis stop.

The simplest interpretation of these data is that one or the other topoisomerase can function as a swivel to allow replication fork movement. Without such a swivel point, the continued unwinding of DNA strands during DNA synthesis quickly causes positive superhelical turns to build up in front of the fork. Such torsional strain is expected to stop fork movement rapidly. It should be emphasized that the two topoisomerases do not necessarily relieve torsional strain in the same way. Topoisomerase II may act behind the replication fork, passing replicated double-strand DNA segments through each other to disentangle them, whereas topoisomerase I acts by nicking and closing DNA ahead of the fork⁸.

We also tested whether the double mutant was defective in RNA synthesis. The rate of RNA synthesis was compared for a wild-type strain, the topoisomerase single mutants and the double mutant. Figure 2 shows that all four strains show a rapid decrease in the rate of RNA synthesis after a shift to 37 °C. This inhibition is transient and due to heat shock. The wild-type strain and each of the single *top* mutants recover from heat shock in about one hour and resume normal RNA synthesis. After two hours at 37 °C the *top2 ts* mutant shows another reduction in the rate of RNA synthesis. This occurs at the time cells begin to die as they attempt to segregate their intertwined chromosomes at mitosis^{2,3}. The *top1 top2* double mutant, on the other hand, never resumes a normal rate of RNA synthesis, levelling off at 10% of the initial rate (Fig. 2).

As an initial attempt to discover which transcription units were most affected in the double mutant, RNA pulse-labelled with ³H-uracil was fractionated into poly(A)⁻ and poly(A)⁺

* Present address: Department of Biochemistry and Biophysics, University of California Medical Center, San Francisco, California 94143 USA.

† To whom correspondence should be addressed.

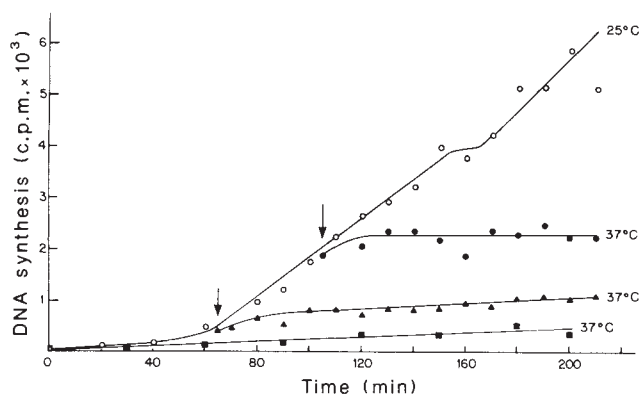


Fig. 1 DNA synthesis in a *top1 top2* mutant. Strain SD119 *MATa top1-1 top2-1* was grown at 25°C, arrested with α -factor, and released from α -factor arrest into medium containing [5, 6-³H]uracil at 25°C (○) or 37°C (■) at zero time. Arrows, times at which portions of the 25°C culture were shifted to 37°C (●) or 25°C (▲). The amount of ³H-uracil incorporated into DNA was measured by determining the amount of alkali-resistant, acid-precipitable counts at the indicated times.

Methods. Cells were grown at 25°C in Y minimal medium¹⁰ supplemented with glucose, yeast extract and amino acids to a density of 7×10^6 cells ml⁻¹. The pH of the medium was then lowered to 4.5 with HCl, and yeast α -factor (Sigma) was added to a final concentration of 10 μ g ml⁻¹. Cells were held in this medium for 1.5 doubling times, pelleted, washed once with one volume Y medium, and resuspended in fresh Y medium containing 20 μ Ci ml⁻¹ [5, 6-³H]uracil (ICN). The culture was split, and half was placed at 25°C and half at 37°C. Portions of the 25°C culture were shifted to 37°C at the beginning or half-way through S phase, the onset of which was judged by the appearance of small buds. At various times 0.2-ml aliquots were removed and placed in 0.5 ml of a stop solution (15% trichloroacetic acid, 200 μ g ml⁻¹ thymine and 50 mM sodium pyrophosphate) containing 0.1 ml of unlabelled stationary-phase yeast cells as carrier. The samples were pelleted, resuspended in 50 μ l of 0.6 M NaOH, and placed at 37°C overnight for RNA hydrolysis. Samples were neutralized with 6 M HCl and precipitated with 1 ml stop solution on ice for 1 h. The precipitates were collected on 0.45- μ m nitrocellulose filters, and the filters washed, dried, placed in toluene-Omnifluor (NEN), and radioactivity determined in a scintillation counter.

Fig. 3 Fluorogram of pulse-labelled RNA in *TOP*⁺ and *top1 top2* strains. Cultures growing at 25°C were split in half at zero time, one half kept at 25°C and the other shifted to 37°C. At various times (indicated in minutes after time zero above each lane of the fluorogram) aliquots were pulse-labelled at the appropriate temperature with [5, 6-³H]uracil. Total RNA was isolated and separated on composite agarose-acrylamide gels and processed for fluorography. *a*, Wild-type (*TOP*⁺); *b*, *top1 top2* mutant; *c*, same fluorogram as *b* but the bottom portion after an 8-fold longer exposure. The positions of 5S and 4S RNA were verified by electrophoresis of purified ³²P-labelled 5S RNA and tRNA, a gift from J. Andersen and N. Delihias.

Methods. Cells were grown and labelled as in Fig. 2 except that 3-ml aliquots were labelled with 30 μ Ci ml⁻¹ [5, 6-³H]uracil for either 10 min (SD118 *TOP*⁺) or 30 min (SD119 *top1-1 top2-1*). The reaction was stopped by chilling with 3 vols of ice-cold sterile water. The sample was pelleted, washed in 3 vols of ice-cold sterile water, and the pellet frozen for later RNA isolation. RNA was isolated as described¹¹ and separated on a 2.5% acrylamide/0.5% agarose gel¹². Processing for fluorography was as described¹³.

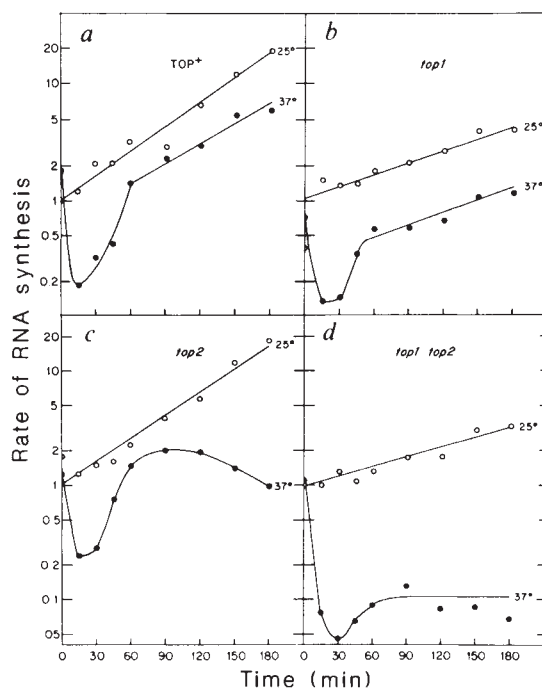
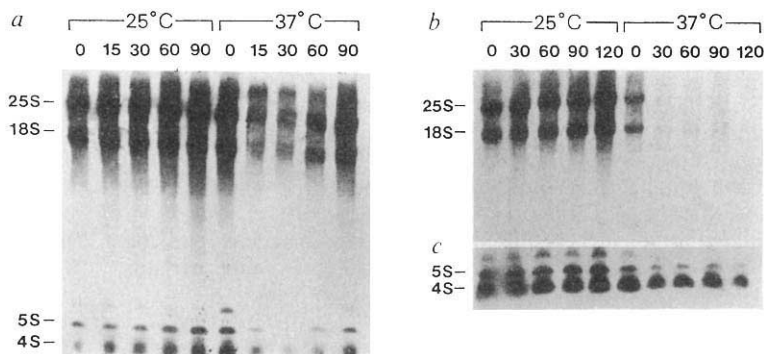


Fig. 2 Rates of RNA synthesis in topoisomerase mutants. Cultures growing at 25°C were split in half at zero time, one half kept at 25°C and the other shifted to 37°C. At intervals aliquots were pulse-labelled with [5, 6-³H]uracil at the appropriate temperature for 5 min. The rate of RNA synthesis is the total acid-precipitable radioactivity incorporated in the 5 min pulse; ○, 25°C; ●, 37°C. Rates normalized relative to 25°C zero time value for each strain. **Methods.** Cells were grown at 25°C in Y medium (as in Fig. 1) to a density of 2×10^7 cells ml⁻¹. The culture was then split into two parts, one at 25°C and the other at 37°C. At various times 1 ml of each culture was removed and incubated with 5 μ Ci [5, 6-³H]uracil for 5 min at the appropriate temperature. The reaction was stopped by the addition of 2 ml stop solution as in Fig. 1. The sample was pelleted and the pellet resuspended in 2 ml fresh 15% stop solution and kept on ice for 30 min. The sample was then filtered and radioactivity determined as in Fig. 1. Strains used were SD118 *TOP*⁺, SD116 *top1-1*, SD117 *top2-1* and SD119 *top1-1 top2-1*. Similar results were obtained with strains containing the *top1-7* and *top1-8* null mutations⁶.



pools on an oligo(dT)-cellulose column. The poly(A)⁺ pool contains mainly RNA polymerase II transcripts and the poly(A)⁻ pool contains RNA polymerase I and III transcripts (ribosomal RNA and tRNA). The rates of RNA synthesis for these two types of RNAs in the double mutant were compared. The rate of poly(A)⁻ RNA synthesis is more severely affected than is poly(A)⁺ RNA synthesis. For example, after 100 min at

37°C, the rate of poly(A)⁻ RNA synthesis is 15%, and the rate of poly(A)⁺ RNA synthesis is about 30% of their respective initial rates.

These results suggested that rRNA and/or tRNA synthesis might be most affected in the double mutant. The synthesis of these transcripts was specifically analysed by pulse-labelling RNA with ³H-uracil and subjecting the RNA to gel elec-

trophoresis and visualizing it by fluorography. Wild-type cells pulse-labelled at 37 °C show a transient decrease in the rate of synthesis for all three rRNA species (25S, 18S and 5S) as well as tRNA (4S) (Fig. 3a). This is followed by partial recovery from heat shock by 60 min and full recovery by 90 min. In contrast, double mutant cells show a rapid and drastic decrease in the synthesis of both 25S and 18S rRNAs, with no recovery even after 120 min at 37 °C (Fig. 3b). The synthesis of 5S rRNA appears to be less affected than that of the large rRNAs, and tRNA synthesis seems to be relatively unaffected. These impressions are confirmed by cutting out the RNA bands from the fluorograms and determining the amount of radioactivity in each band. After 90 min at 37 °C the rate of 25S and 18S rRNA synthesis is 10% of the initial rate, but the rates of 5S rRNA synthesis and tRNA synthesis are 24% and 93% of their initial rates, respectively. This confirms that RNA polymerase I transcription is most affected by inactivating topoisomerase activity.

Although rRNA synthesis is most affected in the topoisomerase double mutant, the synthesis of poly(A)⁺ RNA also is inhibited to some extent. We therefore examined the effect on a specific RNA polymerase II transcription unit, the tightly regulated *GAL1* gene, which codes for galactokinase. There is very little transcription of this gene in the absence of galactose and high expression after galactose induction. Wild-type and *top1 top2* mutant cells were grown at 25 °C in a non-inducing carbon source, shifted to 37 °C and induced with galactose 15 min after the temperature shift. Control cultures were induced at 25 °C. At various times after induction, samples were removed and assayed for galactokinase, as a measure of *GAL1* expression. The wild-type strain has a 2.5-fold greater rate of induction than does the double mutant, but this difference is seen at both 25 °C and 37 °C, presumably because the wild-type strain grows more rapidly. There is no hint of slower induction of *GAL1* in the double mutant at 37 °C compared to 25 °C; we therefore conclude that even after the inactivation of both topoisomerases, efficient synthesis of *GAL1* mRNA can take place. The synthesis of two other messenger RNAs, the products of the *MATa1* and the *SIR3* genes, was also found to be unaffected in the *top1 top2* double mutant (data not shown).

We have presented evidence that DNA topoisomerases are required for DNA replication and rRNA transcription in yeast. Surprisingly, topoisomerase I and topoisomerase II can substitute for each other in these processes. Saavedra and Huberman have shown that either enzyme can also substitute for the other in the *in vivo* relaxation of torsionally strained DNA of the 2-micron plasmid⁹. The only essential function known in which one topoisomerase cannot substitute for the other is in the segregation of sister chromatids at mitosis. This is probably because only topoisomerase II is able to decatenate intact DNA molecules. On the other hand, both enzymes can efficiently relax negatively and positively supercoiled DNA. We view their ability to substitute for one another as a reflection of this similarity. Accordingly, both enzymes seem to be capable of acting as a swivel, using different strand-passing mechanisms to relax the torsional strain that arises during replication, and, perhaps surprisingly, during transcription at the rDNA loci.

Very recently it has been reported that a fission yeast *top1 top2* ts mutant is also defective in DNA and RNA synthesis⁵. Thus, the biological functions of topoisomerases in two very different yeast species appear to be similar. It is not clear whether the situation in yeast can be generalized to higher eukaryotes. The functions that have been uncovered for the topoisomerases in yeast should be conserved, but it is possible that additional or more specialized roles have been added during the evolution of multicellular organisms.

We thank P. J. G. Butler for an important suggestion, Allan Miller for help with the experiment on *MATa1* and *SIR3* mRNAs, and Bruce Alberts and Keith Yamamoto for their comments on the manuscript. This work was supported by NIH grant GM28220.

Received 7 October 1986; accepted 14 January 1987.

1. Thrash, C., Voelkel, K., DiNardo, S. & Sternglanz, R. *J. biol. Chem.* **259**, 1375-1377 (1984).
2. DiNardo, S., Voelkel, K. & Sternglanz, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2616-2620 (1984).
3. Holm, C., Goto, T., Wang, J. C. & Botstein, D. *Cell* **41**, 553-563 (1985).
4. Uemura, T. & Yanagida, M. *EMBO J.* **3**, 1737-1744 (1984).
5. Uemura, T. & Yanagida, M. *EMBO J.* **5**, 1003-1010 (1986).
6. Thrash, C., Bankier, A. T., Barrell, B. G. & Sternglanz, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4374-4378 (1985).
7. Goto, T. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7178-7182 (1985).
8. Champoux, J. J. & Been, M. D. in *Mechanistic Studies of DNA Replication and Genetic Recombination* (ed. Alberts, B. M.) 809-815 (Academic, New York, 1980).
9. Saavedra, R. A. & Huberman, J. A. *Cell* **45**, 65-70 (1986).
10. Zakian, V. A., Brewer, B. J. & Fangman, W. L. *Cell* **17**, 923-934 (1979).
11. Elion, E. A. & Warner, J. R. *Cell* **39**, 663-673 (1984).
12. Peacock, A. C. & Dingman, C. W. *Biochemistry* **7**, 668-674 (1968).
13. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83-88 (1974).

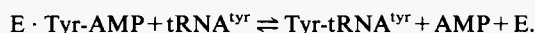
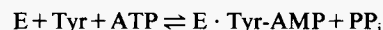
Structure of a mutant of tyrosyl-tRNA synthetase with enhanced catalytic properties

K. A. Brown, P. Brick & D. M. Blow*

Blackett Laboratory, Imperial College, London SW7 2BZ, UK

One surprising outcome of applying the techniques of protein engineering to the enzyme tyrosyl-transfer RNA synthetase was that the enzyme's activity could actually be increased by a specific sequence change¹. In particular, altering residue threonine 51 to a proline (mutant TP51) increased the enzyme's affinity for tyrosyl adenylate complexes². The non-additive effect of combining the TP51 mutation with a second sequence alteration (histidine 48 to a glycine)³ suggested that the effect of the TP51 change might be mediated by a structural change involving the peptide backbone^{1,3}. To address the question of the mechanism by which the TP51 change increases the activity of tyrosyl-tRNA synthetase we have determined the structure of the mutant enzyme. We find the change has a purely local effect on the structure of the enzyme, and conclude that the increased activity of the TP51 mutant probably results from the replacement of the polar threonine residue by a non-polar group: in the wild-type enzyme substrate binding is disfavoured by the displacement of solvent from the vicinity of threonine 51. This unfavourable effect is absent in the TP51 mutant.

The tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* catalyses the aminoacylation of its tRNA by the following two-step reaction⁴:



The TP51 mutant shows a 15-fold decrease in K_M (ATP), which, with a small increase in k_{cat} , leads to a 25-fold increase in k_{cat}/K_M for the pyrophosphate exchange reaction¹. Mutants TA51 and TC51, containing alanine and cysteine at this position, show intermediate kinetic properties but mutants TG51 and TS51, containing glycine and serine respectively, show much poorer activity compared to wild-type⁵. In a more detailed study² the free energies of the enzyme-bound intermediates were also determined. Comparison of the free energy profiles for the TP51 mutant and the wild-type shows that all complexes involving bound tyrosyl adenylate are more stable in the mutant by at least 2.2 kcal mol⁻¹.

The enzyme exists mainly as the tyrosyl adenylate complex under physiological conditions⁶. *In vivo* the rate-limiting step is the aminoacylation of tRNA. Ho and Fersht² point out that the TP51 mutant, though catalytically efficient in forming the tyrosyl adenylate complex, would be less efficient *in vivo* in forming tyrosyl-tRNA because the enzyme-bound adenylate complex constitutes a free energy minimum.

* To whom correspondence should be addressed.

Fig. 1 *a*, Main chain atoms of wild-type tyrosyl-tRNA synthetase, residues 46–54, with the adenosine moiety of the bound adenylate substrate shown for clarity. The side chain of Thr 51 is shown. The side chain of Pro 51 as observed from the electron density of the TP51 mutant is indicated in thinner bonds. The dashed lines indicate close polar interactions which are stereochemically favourable for hydrogen bonds in the refined structure. Beginning with the CO48...NH52 interaction, regular α -helical hydrogen bonds are made, but NH51 makes no hydrogen bond to the main chain, and the CO46...NH50 interaction is through a well-ordered water molecule, indicated as two concentric circles. *b*, One section from the electron density difference map between the TP51 crystals and the wild-type obtained using Fourier coefficients ($F_{TP51} - F_{wildtype}$) and phases calculated from the refined wild-type structure including the well-ordered solvent molecules. Contours at intervals of 1.8 standard deviations, with zero contour omitted and negative contours dashed. The strongest features represent the substitution of proline for threonine. Other significant features (above and left of threonine) represent solvent rearrangements. Scale, 2 mm per Å. *c*, Stereo view of the difference density close to Thr 51, contoured at intervals of 1.8 standard deviations. Near Thr 51, positive density is observed at the C γ and C δ positions for proline, with an equal amount of negative density close to the positions of O γ 1 and C γ 2 of the original threonine. The negative density above Thr 51 and the positive density to the left probably represent a rearrangement of solvent which is displaced from its position in the wild-type structure. Both these features exceed 3.6 standard deviations. The negative peak is within hydrogen bonding distance of O47 of the wild-type enzyme. These differences are below the level of a well-ordered water molecule, and no specific solvent sites are observed in the vicinity of the polar groups O47, N51 or O51 in the wild-type structure. Scale, 4 mm per Å.

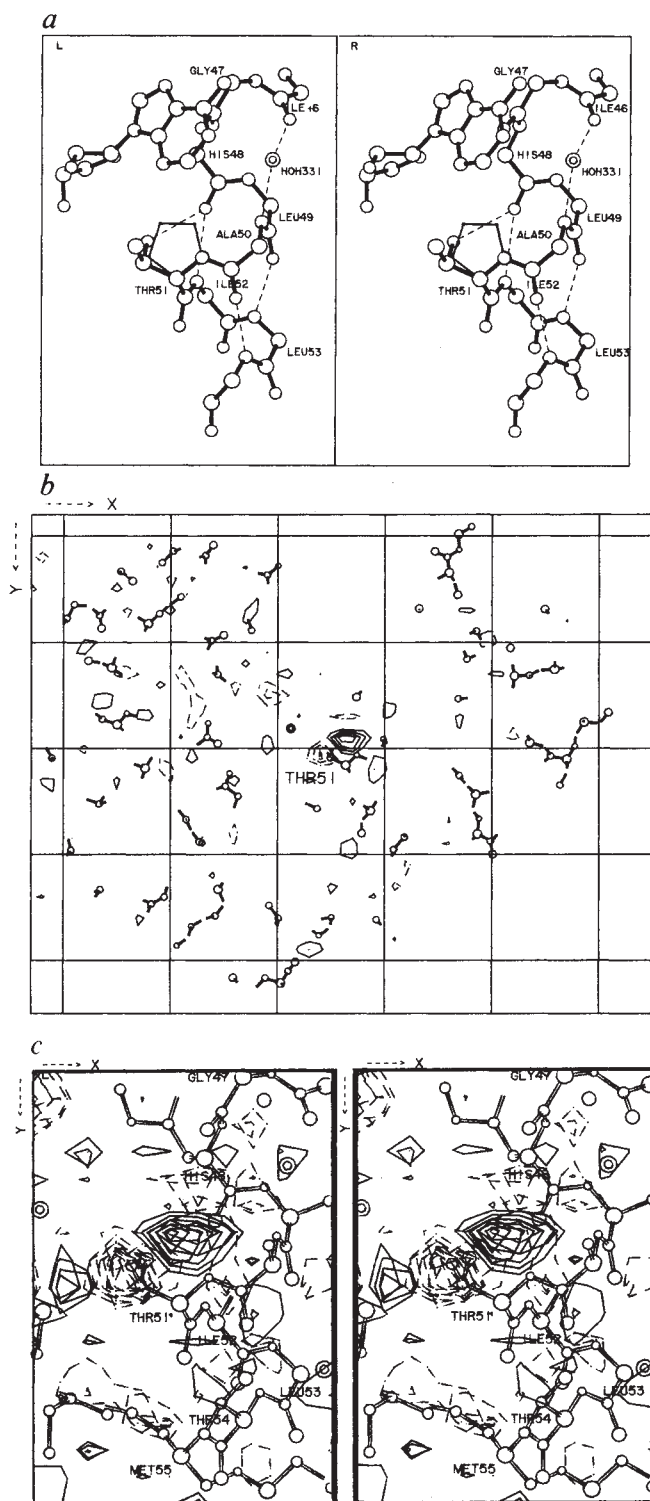
Several mutations of the non-essential His 48 residue have also been studied including HG48 (glycine) and the double mutant HG48TP51 (refs 3, 7). In both HG48 and the double mutant, the free energy barrier at the transition state is about 1 kcal mol $^{-1}$ greater than in the wild-type. Thus, the free energy changes contributed by the two separate mutations, HG48 and TP51, are not additive. This was originally interpreted as evidence of a structural distortion induced by the TP51 mutation, propagating its effect to the mutation at residue 48 (ref. 3).

A more detailed interpretation of all these observations has been gained from a crystal structure analysis of the TP51 enzyme. The enzyme was purified and crystallized from ammonium sulphate under conditions similar to those for the wild-type enzyme⁸.

In our preliminary unrefined structure of *B. stearothermophilus* tyrosyl-tRNA synthetase⁹, residues 46–60 were assigned as α -helical. A more detailed study, confirmed later by refinement (P.B., D.M.B. and T. N. Bhat, in preparation) shows the amino terminus of this helix to be irregular. Although the conformational angles all lie in the α -helical region of the (ϕ , ψ) plot, there is no hydrogen bond from NH51 to CO47. This conclusion is confirmed by the structure determination of a truncated mutant¹⁰, in which 100 amino acids are removed from the carboxyl terminus¹¹, which has allowed the wild-type structure determination to be improved (P.B., D.M.B. and T. N. Bhat, manuscript in preparation).

The electron-density difference map is shown in Fig. 1, together with a presentation of the refined wild-type structure for the adjacent residues. The figure shows how the features of the difference map can be precisely accounted for, without any consequential structural change of the peptide backbone. The proline side chain can fit into the distorted helical structure at the amino terminus of the α -helix, without movement of any other side chains.

In the ligand-free wild-type enzyme, the polar hydroxyl group of Thr 51 is at hydrogen-bonding distance from O48, but is also exposed to solvent. This hydroxyl group approaches a polar group (the ribose O4') when the tyrosyl adenylate intermediate is bound. This interaction had been considered a possible candidate for a hydrogen bond¹. Recent refinement shows, however,



that the distance from this hydroxyl to O4' is about 3.6 Å¹⁰. There is thus no direct hydrogen bond between them.

In the tyrosyl adenylate complex¹², the adenine ring lies close to the Thr 51 side chain and is in van der Waals' contact with the C α and carbonyl oxygen atoms of Gly 47. It encloses a small vacant void, too small to accommodate a water molecule, adjacent to the side chain of the Thr 51 (see Fig. 2a). In the ligand-free form of the enzyme, partially ordered solvent (WAT-330)¹⁰ occupies this void and is displaced upon binding of the adenylate.

In addition, solvent accessibility calculations (refs 13 and 14, and K. S. C. Reid, personal communication) were carried out

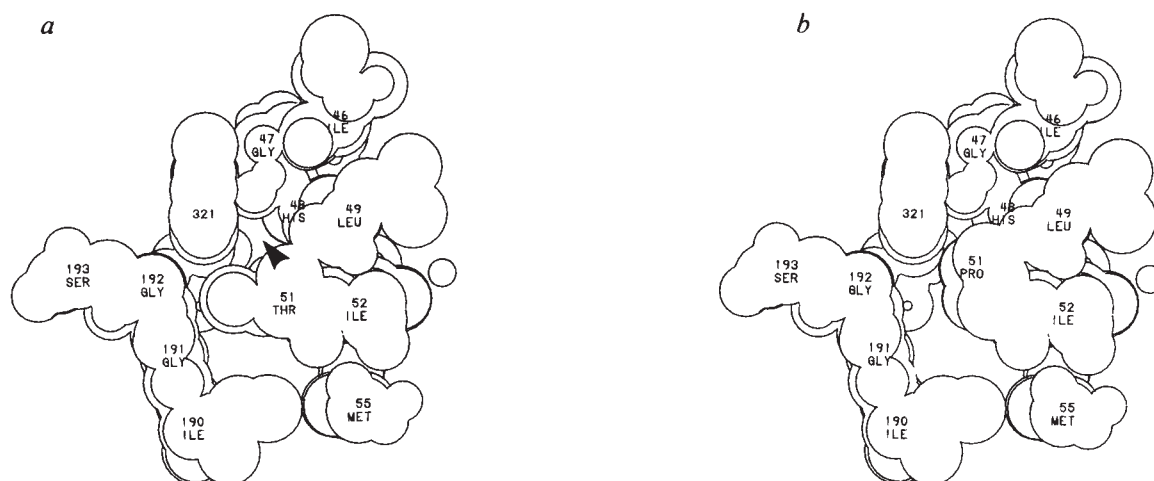


Fig. 2 *a*, Slices through van der Waals' envelopes containing residues 46–55, 190–193, and the adenosine moiety of the adenylate labelled 321. The 'void', indicated by an arrow, is the empty volume enclosed by the side chain of Thr 51 and main chain components from Leu 49, His 48 and Gly 47. Note the proximity of the CO47 to the centre of the adenine ring. In the unbound form of the enzyme this carbonyl group can form a hydrogen to WAT-330 which is displaced in the TP51 mutant. *b*, The same as *a* except with proline substituted for threonine at position 51. In this mutant form of the synthetase, the C γ and the C δ of the pyrrolidine ring occlude a large percentage of the original volume of the void. (Pictures produced by a computer program¹⁸ by A. M. Lesk and K. D. Hartman.)

on those mutants which primarily affect the binding of adenylate complexes and compared with those for the ligand-free wild-type enzyme. In the TA51 mutant, the polar hydroxyl group of the Thr 51 side chain is lost and the void is slightly more open to bulk solvent. In the TP51 mutant, the void has a greatly reduced volume (see Fig. 2*b*) and the local hydrophobicity of the mutated region has increased. In the TG51 and the TS51 mutants, the void is larger and is accompanied by the loss of the hydrophobic methyl group of the Thr 51 side chain.

Figure 1*c* presents a more detailed view of the electron density differences close to Pro 51 using a three-dimensional presentation. In addition to the major changes, Fig. 1*c* shows other positive and negative features near the Thr 51 side chain which are well above the general noise level of the map, but less than a fully occupied atom site. These features all appear outside the protein density, in solvent regions. In particular, the C γ and C δ of the pyrrolidine ring appear to displace solvent (WAT-330) from the void described in the previous paragraph. The increased hydrophobicity of this region and improved van der Waals' contacts would be expected to improve the enzyme's affinity for tyrosyl adenylate complexes.

Kinetic studies indicate that there is a favourable interaction between the tyrosyl adenylate complex and the side chain of His 48 (refs 3, 7, 15 and 16) although current structural data cannot verify this¹⁰. In the HG48 mutant this interaction is abolished leading to an increased transition state energy^{3,16}. In the HN48 mutant the asparagine side chain appears to mimic the polar interaction of the original histidine with the adenylate and its transition state energy is only slightly increased. But in the HQ48 mutant the longer glutamine side chain may sterically hinder proper binding of the adenylate, as evidenced by an increase in its transition state energy⁷. In the double mutants where proline is also substituted at position 51, the energies of interaction of these side chains with the transition state for the activation of tyrosine are not additive^{3,7}.

The difference map shows that these effects in the double mutants are probably not caused by a direct structural change communicated from one side chain to the other. The cause is a more subtle one, probably due to solvent reorganization. The effect of removing the histidine at side chain 48 and replacing it with glycine is to increase the solvent-accessible surface for residue 51 as well as other neighbouring residues exposed in this region in the ligand-free enzyme. If we choose to interpret a change in solvation as a change in structure, then indeed this

may account for the non-additivity of the free energy changes due to mutations at residues 48 and 51 (ref. 3).

Crystallization and structural studies are in progress on some of the mutants mentioned here besides TP51. In particular, we hope to examine other mutations at position 51 for evidence of solvent reorganization which may correlate with the predicted changes in the accessibilities of differing residues. We also hope to study the changes at His 48 to verify that no structural change has occurred and to check for the predicted changes in solvation in the region previously described. We must also verify that the orientation of the tyrosyl adenylate in the active site does not significantly change.

We have learned from this study that one must allow for changes in surrounding solvent as well as changes in main chain structure and side chain character when trying to interpret the catalytic effects of engineered mutations. Admittedly, fine details of solvent organization are not easy to observe in electron-density difference maps and their kinetic consequences are difficult to compute¹⁷. But prediction of the effects of future factitious mutations must attempt to incorporate solvation changes if one hopes that protein engineering will become less of an empirical activity.

K.A.B. was a Fulbright Scholar and is a recipient of an ORS award from the UK Committee of Vice-Chancellors and Principals. This work is supported by the UK Medical Research Council.

Received 19 November 1986; accepted 16 January 1987.

1. Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, P. & Winter, G. *Nature* **307**, 187–188 (1984).
2. Ho, C. K. & Fersht, A. R. *Biochemistry* **25**, 1891–1897 (1986).
3. Carter, P. J., Winter, G., Wilkinson, A. J. & Fersht, A. R. *Cell* **38**, 835–840 (1984).
4. Fersht, A. R. & Jakes, R. *Biochemistry* **14**, 3350–3356 (1975).
5. Fersht, A. R., Wilkinson, A. J., Carter, P. & Winter, G. *Biochemistry* **24**, 5858–5861 (1985).
6. Mulvey, R. S. & Fersht, A. R. *Biochemistry* **16**, 4731–4737 (1977).
7. Lowe, D. M., Fersht, A. R., Wilkinson, A. J., Carter, P. & Winter, G. *Biochemistry* **24**, 5106–5109 (1985).
8. Reid, B. R., Koch, G. L. E., Boulanger, Y., Hartley, B. S. & Blow, D. M. *J. molec. Biol.* **80**, 199–201 (1973).
9. Bhat, T. N., Blow, D. M., Brick, P. & Nyborg, J. *J. molec. Biol.* **158**, 699–709 (1982).
10. Brick, P. & Blow, D. M. *J. molec. Biol.* **194**, 287–298 (1987).
11. Wayne, M. M. Y., Winter, G., Wilkinson, A. J. & Fersht, A. R. *EMBO J* **2**, 1827–1829 (1983).
12. Monteilhet, C. & Blow, D. M. *J. molec. Biol.* **122**, 407–417 (1978).
13. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
14. Connolly, M. L. *Science* **221**, 709–713 (1983).
15. Fersht, A. R. *et al. Nature* **314**, 235–238 (1985).
16. Wells, T. N. C. & Fersht, A. R. *Biochemistry* **25**, 1881–1886 (1986).
17. Finney, J. L., Goodfellow, J. M., Howell, P. L. & Vovelle, F. *J. biol. Macromol.* **3**, 599–622 (1985).
18. Lesk, A. M. & Hardman, K. D. *Science* **216**, 539–540 (1982).

High-speed countercurrent chromatography

Y. Ito

Support-free high-speed countercurrent liquid chromatography provides a rich domain of applications, some beyond reach of conventional liquid chromatography.

COUNTERCURRENT chromatography or CCC is a generic term for support-free partition chromatographic methods. CCC eliminates complications caused by solid supports, such as adsorptive sample loss and deactivation, tailing of solute peaks, and contamination. However, the lack of a solid support matrix requires the combined use of a unique column geometry and an applied force field to retain the stationary phase in the column against a continuous mobile phase flow, while providing efficient mixing between the two solvent phases to promote partition. Consequently, the existing CCC instruments have taken diverse forms and make use of mechanisms which are quite different from those in conventional liquid chromatography. Since the advent of CCC in 1970¹, partition efficiency of the method has been remarkably improved by shortening separation times from days in droplet CCC² to only a few hours in high-speed CCC³.

Planetary motion

High-speed CCC utilizes the intriguing hydrodynamic behaviour of two immiscible solvent phases in a coil undergoing a particular mode of planetary motion generated by a synchronous coil planet centrifuge. The centrifuge holds a spool-shaped coil holder which revolves around the central axis of the centrifuge, and simultaneously rotates about its own axis at the same angular velocity in the same direction. In doing so, the holder maintains its axis parallel to, and at a fixed distance from, the central axis of the centrifuge. This synchronous planetary motion fulfills two seemingly magical tasks, both essential for performing high-speed CCC. First, the synchronous rotation of the holder constantly unwinds the twist of the flow tubes caused by revolution. This permits continuous elution through the rotating coil without the use of a conventional rotary seal device, which can be a potential source of leakage and contamination. The second task is even more important. When the coiled column is coaxially mounted around the coil holder, the planetary motion of the holder unilaterally distributes the two solvent phases in the column in such a way that one phase (the head phase) entirely occupies the head side, and the other phase (the tail phase) occupies the tail side of the coil.

The head-tail relationship refers to the archimedean screw force acting on the

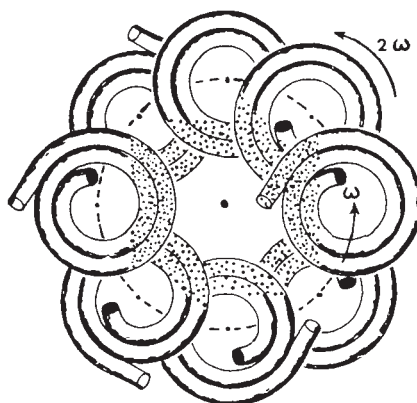


Fig. 1 Mixing and settling zones of a spiral column in a high-speed CCC centrifuge. The head corresponds to the internal terminus and the tail to the external terminus.

rotating coil, where all objects of different density are driven from the tail toward the head of the coil. This hydrodynamic phenomenon can be effectively utilized for performing CCC in the following two ways. The coil can be entirely filled with the tail phase followed by elution with the head phase from the tail toward the head. Alternatively, the coil can be filled first with the head phase followed by elution with the tail phase from the head toward the tail. In either case the hydrodynamic phenomenon facilitates rapid movement of the mobile phase through the stationary phase, yielding extremely high retention of the stationary phase in the coil.

The system also permits simultaneous introduction of the two solvent phases through the respective ends of the coil — the head phase from the tail and the tail phase from the head. This dual countercurrent operation requires an additional flow channel at each end of the coil to collect the effluent and a sample feed

line at the middle portion of the coil. This dual CCC system is unique to high-speed CCC.

The high partition efficiency inherent to high-speed CCC is largely attributed to a peculiar solvent interaction in the rotating coil, as observed by W.D. Conway in our laboratory⁴. Figure 1 illustrates successive positions of a spiral column undergoing planetary motion as indicated by arrows. The area of the spiral column is divided into two distinct zones: the mixing zone with evidence of vigorous agitation of the two solvent phases, and the settling zone showing clear separation of the two layers, with the heavier phase in the outer portion of each spiral segment and the lighter phase in the inner portion. The mixing zone is always located near the centre of the centrifuge while the column rotates around its axis twice during each revolution. This fact clearly indicates that each mixing zone is moving toward the head of the spiral column at a rate of one turn per revolution. When the column is rotated at 800 r.p.m., the two solvent phases are subjected to a repetitive mixing and settling process at the enormously high frequency of over 13 times per second, while the mobile phase is continuously passing through the stationary phase. The above finding promises a high partition efficiency for the system, even with a high mobile phase flow rate.

Applications

Application of high-speed CCC may be classified into three categories: continuous extraction with a short coil, chromatographic separation with a multilayer coil, and dual CCC. Reports on continuous extraction are so far limited to successful enrichment of urinary daunorubicin anti-

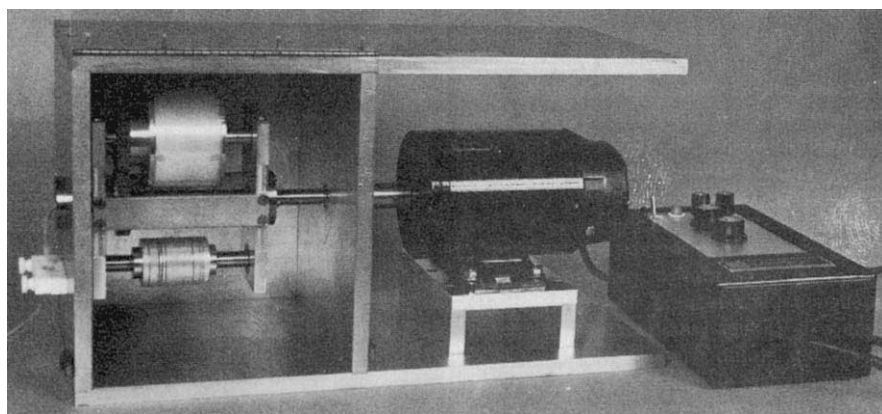


Fig. 2 An analytical high-speed CCC centrifuge.

tumour drug metabolites⁵. On the other hand, high-speed CCC has been used in the semipreparative chromatography of a variety of biological compounds including antibiotics, steroids and other drugs, various agrochemicals, dyes, tannins, and natural and synthetic peptides, as summarized elsewhere³. The majority of these applications have been carried out with the Ito multilayer coil separator and extractor (I have no affiliation with the company) manufactured by P.C. Inc., in Potomac, Maryland. Recently, Pharma-Tech Research Corporation in Baltimore, Maryland began marketing several models covering preparative to analytical separations. A few more companies are expected to join the marketing race in the near future.

Further refinements

Current refinements of the dual CCC technique include foam CCC and liquid-liquid dual CCC. Although these methods are still in the developmental stage, the potential capability of foam CCC has been successfully demonstrated in separations of test samples of dyes, plant hormones, and blood proteins^{6,7}. Efficient chromatographic separations of nine steroids on a liquid-liquid dual CCC system have been reported by Lee *et al.*⁸.

Future developmental prospects include interfacing high-speed CCC with infrared and mass spectroscopy techniques. High-speed CCC/FT-IR spectroscopy has been pioneered by Romanach and de Haseth at the University of Georgia, Athens, with promising results⁹. Figure 2 shows an analytical high-speed CCC centrifuge recently fabricated at the US National Institutes of Health for high-speed CCC/mass spectroscopy using a thermospray method. This apparatus can be operated at 2,000 r.p.m. to yield a partition efficiency of 1,600 theoretical plates for microgram quantities of samples¹⁰. Separations generally require 60 to 90 minutes, but may be substantially shortened by refinement of the centrifuge design. □

Yoichiro Ito is at the Laboratory of Technical Development, National Heart, Lung, and Blood Institute, Bethesda, Maryland 20892, USA. For additional information fill in reader service number 100.

1. Ito, Y. & Bowman, R.L. *Science* **167**, 281-283 (1970).
2. Tanimura, T., Pisano, J.J., Ito, Y. & Bowman, R.L. *Science* **169**, 54-56 (1970).
3. Ito, Y. *CRC Crit. Rev. analyt. Chem.* **17**, 65-143 (1986).
4. Conway, W.D. & Ito, Y. 1984 Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy. Abstr. No. 471 (1984).
5. Nakazawa, H. *et al. J. Chromatogr.* **307**, 323-333 (1984).
6. Ito, Y. *J. Liq. Chromatogr.* **8**, 2131-2152 (1985).
7. Bhatnagar, M. & Ito, Y. 1986 Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy. Abstr. No. 1050 (1986).
8. Lee, Y.W., Cook, C.E. & Ito, Y. 1986 Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy. Abstr. No. 1049 (1986).
9. Romanach, R.J., de Haseth, J.A. & Ito, Y. *J. Liq. Chromatogr.* **8**, 2209-2219 (1985).
10. Ito, Y. & Lee, Y.W. *J. Chromatogr.* **391**, 290-295 (1987).

FASEB highlights

This year's conference of the Federation of American Societies of Experimental Biology on 29 March-3 April in Washington, D.C., will feature products of use to every experimental biologist.

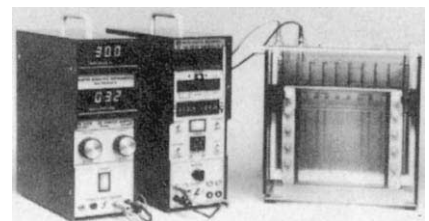
LABCONCO Corporation, in booths 1555, 1557 and 1559, has introduced a 4.5-litre **benchtop freeze drying system**, the Lyph-Lock 4.5 (Reader Service No. 101). The \$3,895 (US) lyophilizer measures only 18 inches wide × 23 inches deep × 22.5 inches high, and saves valuable work space. Its 1/3 horsepower refrigeration system cools the freeze dryer condenser



Labconco's lyophilizer is compact enough for the benchtop.

down to -54°C , and Labconco says the Lyph-Lock 4.5 can remove up to two litres of water every 24 hours. The unit comes equipped with a colour-coded temperature gauge and an electronic vacuum gauge, and has an effluent drain line with a retractable hose located in the front for ease in defrosting. It has an integral ten-port drying chamber, with 0.75 inch and 0.5 inch valve sizes.

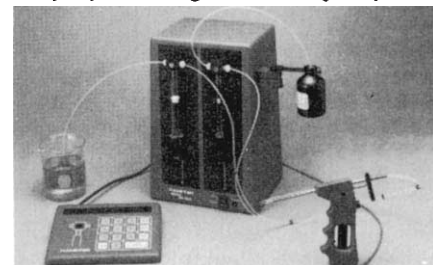
Hoefer Scientific Instruments, in booths 1218, 1220, and 1222, will be showing its PC 750 Pulse Controller for **reverse-field gel electrophoresis** (Reader Service



Hoefer's pulse controller tugs large DNA pieces through a gel.

No. 102). The \$1,800 (US) controller provides periodic reversal of the polarity of the electric field during electrophoresis, making possible the separation of DNA fragments of up to 1,000 kilobases. In addition, smaller DNA fragments can be separated with higher resolution by using the controller to create rapid forward/reverse pulses. The Hoefer Scientific controller provides pulses from 0.1 milliseconds to 999.9 seconds long, and can be connected to any d.c. power supply up to 750 volts. For resolving a broad size range of DNA fragments in one run, the PC 750 has special timing ramps for increasing forward time, reverse time, or a combination of both. With four pairs of output jacks, the PC 750 can run up to four electrophoresis units at once.

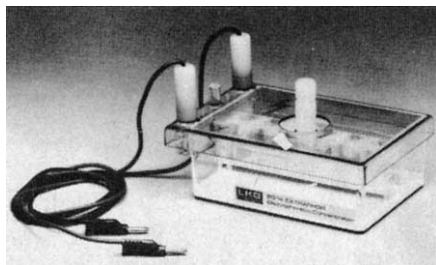
New software is now available for the Hamilton Microlab 900 Series of diluter/dispensers (Reader Service No. 103). Hamilton has developed the controller software to increase programming flexibility by allowing users to specify time



Hamilton's diluters can do more with new software.

delays of up to 99 seconds, and choose either manual or automatic repeat operation. The software has been improved to make the controller easier to program, and includes displays of cumulative volumes to help users keep track of the programming sequence. Hamilton will be in booths 1235, 1237 and 1239.

For **concentrating nucleic acids and proteins** derived from gel samples, LKB Instruments, in booths from 853 to 1062, has the LKB Extraphor electrophoretic concentrator (Reader Service No. 104). The £525 (UK) apparatus can process up to six gel samples simultaneously without the use of dialysis tubes or membranes, that have been known to cause product loss. To use the Extraphor concentrator, gel slices are placed in cup-shaped depressions in one of the chambers and covered with a buffer of low salt molarity, that also flows down through the V-



For recovering substances from electrophoretic gels.

channel and up into the parallel chamber. A syringe is then used to inject 50 μ l of high molarity salt underneath the buffer along the bottom of the V-channel. When current is applied across the two chambers, macromolecules migrate out of the gel slices and enter the V-channel, and can be subsequently drawn off with a pipette. LKB claims separation times of 20–40 min for unfixed and unstained samples, and 60–120 min for fixed and stained samples.

In booths from 1107 to 1314, Beckman will be showcasing all its most recent product offerings. Among the multitude will be the Model 171 **radioisotope detector**, geared toward drug, pesticide and metabolism studies, receptor binding analyses, product purification, hybridization probe desalting, and protein mapping (Reader Service No. 105). The \$12,500 (US) detector is cold room-compatible to enable



An automated sleuth for detecting labelled compounds.

analysis of labile compounds, and provides quantitation of single- and dual-labelled compounds as they are eluted from a liquid chromatography column. With eight independently programmable user files, the radioisotope detector has preset isotope settings for tritium, carbon-14, phosphorus-32, sulphur-35 and iodine-125. It has a cartridge design flow cell that allows quick volume changes and easy decontamination, and standard features include automatic fraction collector control, luminescence correction, automatic background subtraction, automatic start and stop and user-accessed diagnostics. The system is available as a solid system, liquid system, or both.

Plain and twin-frosted **glass microscope slides** are being offered by Elkay, exhibiting in booth 1338 (Reader Service No. 106). The slides come in two thicknesses, 0.8–1.00 mm and 1.0–1.2 mm, with an overall size of 76 mm $\times$ 26 mm. Manu-



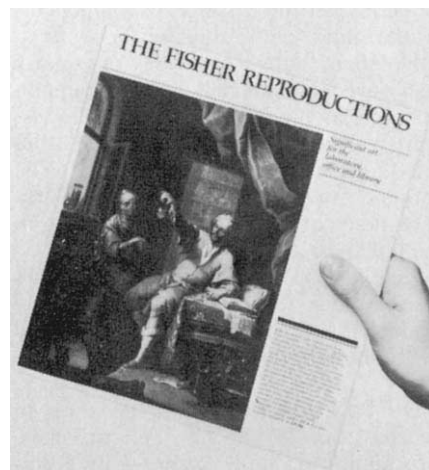
Elkay's array of slides includes plain and twin-frosted.

factured from semi-white glass for optimum clarity, the slides are pre-washed and cleaned ready for use, with all edges ground and polished for safe handling. Slides sell for £24–28 (UK) per thousand in boxes of 50, and free samples are available from the company.

Separation and analysis

The Kratos Division of ABI Analytical has recently introduced two new **generic systems for HPLC** (Reader Service No. 107). The \$19,900 (US) Kratos System 300 is a high-pressure binary gradient system, and at its heart is the Kratos Spectroflow 783 continuously variable UV/VIS absorbance detector. This detector has programmable operation, flow and composition gradient control, and is integral to the system, thus saving bench space. The \$41,000 (US) Kratos System 500 offers more complete automation because it is configured with Kratos' DS 650 chromatography workstation, allowing control of the binary system as well as the autosampler for complete sample verification through bar reading. Both systems feature a pump that uses two pistons with overlapping strokes and a low-volume, integral pulse dampener for pulse-free solvent flow. Applied Biosystems will have an exhibit in booths 1543, 1545, 1547, and 1549.

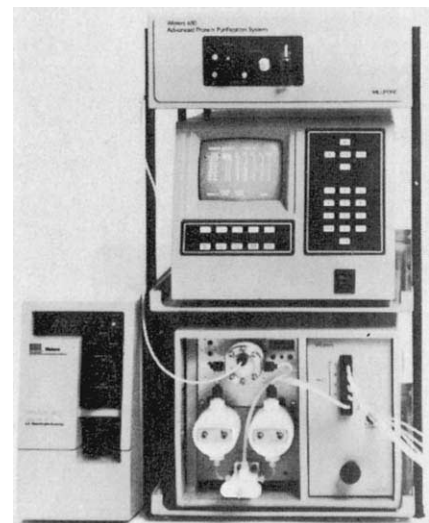
Varian, in booth 1356, has big news with its latest AminoTag **amino acid analyser** (Reader Service No. 108). Varian's machine works by labelling samples with AminoTag (fluorenylmethylchloroformate) in a pre-column derivatization step, yielding stable and highly fluorescent derivatives of both primary and secondary amino acids. Then, automatic liquid/liquid extraction removes excess derivatizing agent, and the derivatized amino acids are separated on a C_{18} reversed phase HPLC column. Resolution is optimized using a ternary solvent gradient, and Varian claims separations can be completed in less than 30 minutes. The AminoTag amino acid analyser in-



FISHER Scientific is sharing its museum collection of art depicting early medicine and science through its **Fisher Reproduction Series** (Reader Service No. 109). The series of prints offers an authentic look at the early days of medicine, biology, metallurgy, chemistry and other disciplines, as drawn and painted by seventeenth and eighteenth century European artists. Twelve reproductions are available, including a five-colour lithograph of "The Physician" by Franz Christophe Janneck (1703-1761), pictured above. The reproductions are available for the cost of printing, framing and shipping alone, and most prints sell for \$30 (US), making them a worthwhile investment in art for the laboratory wall.

cludes a Varian 600 series control and data handling station, which performs comprehensive data manipulation and is designed to present full analysis reports.

In booths from 143 to 242, Waters, a division of Millipore, will be introducing its new **protein purification system** (Reader Service No. 110). The \$14,000 (US) 650 Advanced Protein Purification System features a totally non-metallic fluid path,

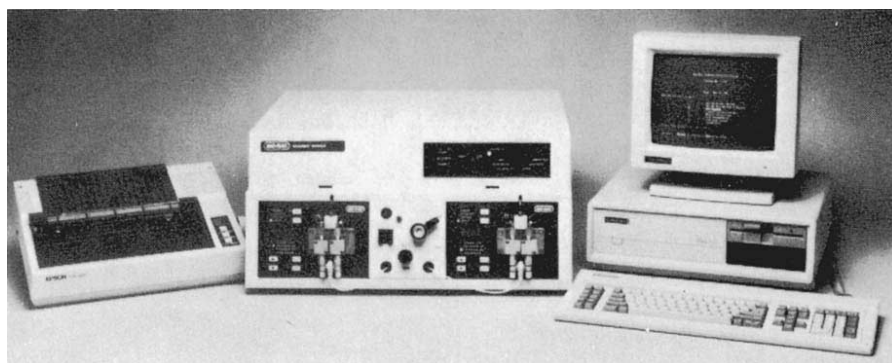


Waters' protein purifier has a non-metallic fluid path.

for better inert processing. All components in the fluid path use fluoropolymers, polyolefins, ruby, sapphire or quartz materials. Columns ranging from 5 mm to 8 cm in diameter may be used on the system, with a choice of flow ranges from 0.1–45 ml min⁻¹ or 0.5–80 ml min⁻¹. The variable wavelength ultraviolet absorbance detector is capable of operating in the 0.005–10 AU range. The Waters 650 System can be used with a range of scalable protein separation chemistries including ion exchange, gel filtration and hydrophobic interaction.

Du Pont will be setting up shop in booths from 735 to 938, and will have information on its BioSeries Poly F column (Reader Service No. 111). This column is designed for **reversed-phase separation of proteins and peptides**, and is packed with a proprietary Du Pont fluoropolymer, that is stable from pH 0 to pH 14, allowing the column to be regenerated with basic eluents. The BioSeries Poly F is a 6.2 mm inner diameter 80 mm long stainless-steel column compatible with all chromatographic systems, and costs \$360 (US). A special introductory price of \$250 is in effect until 1 October 1987.

High-speed countercurrent chromatography is available commercially from P.C., Inc., and will be on display in booth



Bio-Rad's HRLC system may help avoid pressure shock and bed collapse, extending column life.

1662 (Reader Service No. 112). This method uses centrifugal forces to achieve the separation of compounds in a totally liquid system, with no solid stationary phase. An infinite variety of solvents may be used, and no high pressure apparatus is necessary. Countercurrent chromatography has been successfully used to separate compounds such as organic dyes, antibiotics, pesticides, corticosteroids and amino acids. Systems can be configured according to the user's specific needs, and cost \$8,600 (US), including column.

Bio-Rad, in booths 918–923, claims to provide an economical solution to **binary gradient HPLC** and data analysis needs. Its \$13,500–16,000 (US) Series 400 HRLC (high-resolution liquid chromatography)

gradient systems provide binary gradients, up to two channels of data collection and integration, and full system control (Reader Service No. 113). The most salient aspect of the series is the Soft-Start mobile phase pumps that ramp up the flow rate gradually at the initiation of the analysis, avoiding pressure shock and possible bed collapse, and prolonging column life. Two pumps are included in each HRLC system, and each operates over the range from 0.01 ml min⁻¹ to 10 ml min⁻¹. The top of the line Model 402 adds microvolt sensitivity, video strip chart display, data storage, and reintegration to provide complete quantitative analysis.

Promega will have an extensive lineup of reagents, enzymes, and vectors for

ADVERTISEMENTS

VIDEO MICROSCOPY PROBLEMS? MEET THE NEW SOLUTION: DAGE-MTI 70 SERIES

- Interfaces with computer systems
- Compact remote head – 1200 line resolution
- Separate control console with lockable controls
- Automatic or manual operation
- Low linearity distortion (0.25% or less)
- Superior detail contrast

For solving computer interface problems, for digital imaging requirements of any kind, for solving problems with brightfield, darkfield, polarized light, phase – contrast, and D.I.C., the new DAGE-MTI 70 provides compactness, flexibility, and precision performance you've never seen before.

Ask for information on this and other models.

DAGE-MTI, Inc.
208 Wabash Street
Michigan City, IN 46360
Ph. 219/872-5514



Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem

- ☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
Title _____
Institution/Company _____
Department _____
Address _____
City _____ State _____ Zip _____
Phone (____) _____

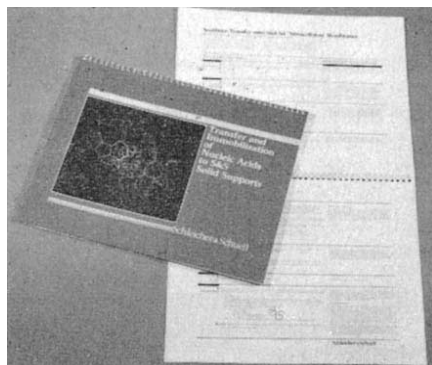
UMI Article Clearinghouse

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

molecular biology in booths 1568 and 1570. Also on display will be Promega's new ProtoBlot **Western blot AP starter systems** (Reader Service No. 114). These \$29 (US) systems employ an alkaline phosphatase based detection system that Promega claims provides greater sensitivity than Protein A or HRP-based systems, and eliminates the need for radioactive isotopes. Second antibody alkaline phosphatase conjugates for the detection of mouse, rabbit, or human primary antibodies are also available. Each system, containing enough reagents for six runs, contains the necessary colour development substrates for optimum sensitivity and maximum signal-to-noise ratios.

Literature and such

A revised and updated version of the Schleicher & Schuell methods chart, *Transfer and Immobilization of Nucleic Acids to S & S Solid Supports* is now available (Reader Service No. 115). The updated version has the latest techniques for Southern and Northern transfers, dot-blot and slot blots, and colony/plaque lifts. Transfer media such as S & S NC nitrocellulose, Nytran positively charged nylon, and Transa-Bind diazotized paper are also described. The free four-colour spiral bound chart contains detailed procedural explanations and an



A good reference for a range of techniques. extensive reference list, and has ample space for user notes. Schleicher & Schuell will be displaying in booths 1119 and 1121.

Issue number six of the *Whatman Chromatography Folio* has been published and is now available free from Whatman, who will be exhibiting in booths 307 and 309 (Reader Service No. 116). This issue answers questions on which reversed phase TLC to use for various applications, and compares four Whatman bonded phases for TLC: ethyl, octyl, octadecyl and diphenyl. The phases were compared by using them to separate over 30 compounds from several classes such as steroids, phenothiazines, sulphonamides and dipeptides. In response to a multitude

of inquiries regarding the separation of organic acids, the folio also includes a section on methods available for the Partisphere HPLC cartridge system.

Accurate Chemical & Scientific Corp., in booths 819 and 821 has just released a catalogue dedicated to **density gradient media** (Reader Service No. 117). The free 12-page publication describes 14 Nycomed density gradient media products for centrifugation, isolation and biological applications. It outlines product features and characteristics, applications and storage information, and includes detailed charts and technical illustrations.

Fisher Scientific will most likely be handing out its new **FisherBiotech mini-catalogue** in booths 1560-1567 (Reader Service No. 118). The catalogue is in a handy folder format, and lists over 200 Fisher reagents and chemicals for biotechnology. It contains sections for amino acids, electrophoresis chemicals, nucleotides, organic and inorganic buffers for media preparation in enzyme reactions, fermentations and tissue culture experiments, and reagents for ultracentrifugation and nucleic acid isolation. A miscellaneous section includes information on surfactants, protein-modification reagents and chromophore substrates.

ADVERTISEMENTS

FUNA[®]-AGAROSE FOR ELECTROPHORESIS

The FUNA[®]-AGAROSE offered by FUNAKOSHI is a highly purified reagent.

Suitable for separating nucleic acids and proteins by electrophoresis and other immunological techniques.

AGAROSE-ME

- * Gel Strength (1.5%) : 1200g/cm²
on or more
- * Gelling temperature : about 36°C
- * Sulfate Content : 0.3% on or less
- * Electroendosmosis : 0.1-0.19

The other grades are available on request.

in small and large quantities.



Write to :

FUNAKOSHI PHARMACEUTICAL CO., LTD.

2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan

Telephone : Tokyo 03-293-2352

Telefax : 81-3-295-5545

Telex : J28489FUNA

Do you want to know about

- toxicity of new artificial sweeteners?
- diet and Alzheimer's disease?
- disposal of hazardous waste?

BIOLOGICAL QUESTIONS

like these
need answers
from CAS ONLINE!

CAS ONLINE. The database of chemistry...and a good place to go if your research concerns biology, pharmacology, biotechnology, and more!

**CAS ONLINE...over 1/3 of our
database is related to life sciences!**

Send to:
Chemical Abstracts Service
Dept. 32886
P.O. Box 3012
Columbus, OH 43210

CAS ONLINE[®]

Available on STN International[®]
The Scientific & Technical Information Network

CAS is a division of the American
Chemical Society

I'd like to find out about
CAS ONLINE and its coverage of
biological information.

Name _____

Title _____

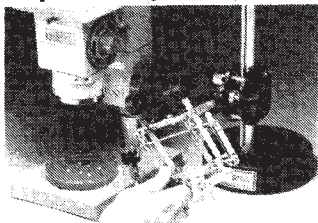
Company _____

Address _____

ADVERTISEMENTS

SINGER INSTRUMENTS

Xenopus oocytes injection?



SINGER MKI MICROMANIPULATOR

We also offer a Microsyringe Kit.

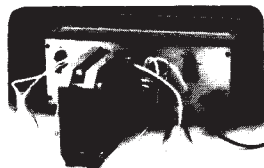
Write to:

Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
 EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIABORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.47

WESTERN BLOTS IN MINUTES

Our new GENIE electrophoretic blotter incorporates space-age plate electrodes for a uniform, high-intensity field. It is easy to load, and uses less than a liter of buffer, yet can blot gels up to 15 by 17 cm. in as little as ten minutes.



only \$190

Idea Scientific Company
P.O. Box 2078
Corvallis, Oregon 97339
(503) 758-0999

Reader Service No.33

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Pharmacia will be in booths from 105 to 208, and will be displaying its new **biotechnology products catalogue** for 1987 (Reader Service No. 119). The catalogue is free from Pharmacia, and includes a new product supplement highlighting Pharmacia's latest developments. Products for chromatography and electrophoresis are also outlined, including the FPLC System and PhastSystem.

Agents and reagents

Amersham claims that its *in vitro* **mutagenesis system** has an 80 per cent efficiency, and is easy to use (Reader Service No. 120). Unlike most mutagenesis techniques which generate heteroduplex mutant molecules, Amersham's system generates homoduplex mutants, and thus avoids the high backgrounds associated with replication of non-mutant strands, and reduces



The instruction manual for Amersham's system.

the time spent screening for mutants. The technique is based on the methodology of F. Eckstein and coworkers, and does not require any specialized vectors or hosts. The \$230 (US) system contains all the enzymes, buffers and special reagents necessary to create a site-specific mutation in a cloned target DNA fragment in combination with a user-supplied mutant oligonucleotide. Amersham will have more information on its mutagenesis system in booths 800–803.

Peptide T, the core peptide within gp120 essential for HIV receptor binding, is now available in analogue form from Bachem, Inc. for AIDS (acquired immune deficiency syndrome) research (Reader Service No. 121). The ([D-Ala¹]-peptide T amide is an octapeptide recently developed by a group at the US National Institutes of Health. It has been shown to be a potent inhibitor of ¹²⁵I-gp120 binding to T4 antigen in brain membranes, and has also been demonstrated to block HIV infection of human T cells at concentrations as low as 0.1 nM. Bachem, exhibiting in booth 420, says the peptide is valuable

in the study of HIV infectivity, and may offer possibilities in the development of neutralizing antibodies and peptide drug therapies for AIDS. The peptide T analogue is available at \$40 (US) for 1 mg, or \$160 for 5 mg.

Boehringer Mannheim Biochemicals will be displaying its newest chemicals and reagents in booths 412 and 414. Perhaps one of the new developments the company is proudest of is the introduction of **NotI endonuclease** (Reader Service No. 122). This enzyme, cleaving sequences at GC/GGCCGC, has proved to be difficult to isolate with a high specific activity in a stable form. Boehringer Mannheim has overcome the isolation obstacles, and is selling its enzyme for £64 (UK) for a 200-unit sample.

A free four-page guide to the design features of the newest **Zeiss microscopes** has been issued by Carl Zeiss, Inc., exhibiting in booths 606, 608, and 610 (Reader Service No. 123). Entitled *Research Microscope Design: New Concepts*, the brochure describes the Zeiss Axioplan and Axiophot — microscopes that allow the user to switch from one technique to another without compromising field of view, resolution or image quality. The booklet describes ICS optics, the Zeiss optical system that permits the insertion of many optical components into the observation path for the full range of microscope techniques, including brightfield, darkfield, phase contrast, qualitative polarized light, differential interference contrast and fluorescence. Also explained is System Integrated (SI) design, a configuration that incorporates permanent or removable optical modules into the microscope housing, permitting quick switching of techniques without impairing image quality or field of view.

Cellular Products, in booth 1144, has introduced two products since the beginning of the year (Reader Service No. 124). The first is a set of murine IgG1 subclass **monoclonal antibodies** (Monobodies) to HIV p17 core protein and gp160/120 envelope core proteins, available in 100 µg vials for \$175 (US). The second is recombinant interleukin-2, manufactured by Cetus Corporation, and prepared in 1 ml fetal bovine serum by Cellular Products. These unglycosylated Monobodies are tissue culture grade, and come at 10,000 BRMP units per millilitre (molecular weight 15.4 × 10⁶). The interleukin-2 antibodies cost \$95 (US) for orders of 1 to 12 vials. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.